2004

A physics-based analytical model of an AlGaN/GaN high electron mobility transistor

Jonathan Sippel

A PHYSICS-BASED ANALYTICAL MODEL OF AN AlGaN/GaN HIGH ELECTRON MOBILITY TRANSISTOR

by

Jonathan C. Sippel

A Thesis submitted in Partial Fulfillment of the

Requirements for the Degree of

MASTERS OF SCIENCE

In

Electrical Engineering

Approved by:

Professor _____
(Dr. Syed Islam – Advisor)

Professor _____
Dr. James Moon – Committee Member)

Professor _____
(Dr. Santosh Kurinec – Committee Member)

Professor _____
(Dr. Robert Bowman – Department Head)

DEPARTMENT OF ELECTRICAL ENGINEERING

COLLEGE OF ENGINEERING

ROCHSTER INSTITUTE OF TECHNOLOGY

RCOHESTER, NEW YORK

MAY 2004

THESIS RELEASE PERMISSION

DEPARTMENT OF ELECTRICAL ENGINEERING

COLLEGE OF ENGINEERING

ROCHESTER INSTITUTE OF TECHNOLOGY

ROCHESTER, NEW YORK

Title of Thesis:

A PHYSICS-BASED ANALYTICAL MODEL OF AN AlGaN/GaN HIGH ELECTRON MOBILITY TRANSISTOR

I, Jonathan C. Sippel, hereby grant permission to Wallace Memorial Library of the Rochester Institute of Technology to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature _____ 5/1/04

Date

ACKNOWLEDGEMENTS

My tenure at RIT has gone by so fast it is hard to believe that I'm approaching my sixth year as a matriculated student in the RIT engineering program. In those six years, I have been enrolled in three engineering programs (luckily ending up in the right one), joined a fraternity, traveled across country, been graced with a second niece, had a brother get married, disjoined from a fraternity, experienced the passing on of my father, and have met numerous people, all of whom have had an influence on my life to some degree. In short, I have experienced a great deal during my time here and would like to thank those who've helped me over the years to grow into the person I am today.

At this time I would like to thank my advisor, Dr. Syed Islam, for his support and guidance throughout this thesis work. He was able to take a scatter brained student, myself, and focus him on a topic he was originally shaky with to grind out a thesis – pretty amazing. If I had to do it all over again, I would still choose to study GaN HEMTs beneath Dr. Islam. Thanks professor.

I would also like to thank Dr. Santosh Kurinec and Dr. James Moon for taking time out of their hectic schedules to be part of my thesis committee. The knowledge and professional insight they have to offer is second to none.

To perform my research, I was outfitted with a computer and desk in the MEMs lab in the EE department alongside fellow researchers Sankha Mukherjee, Brett Klehn, and Shailesh Rai. They played a pivotal role in my retaining any sort of sanity through the research process. Not to mention that every time I felt like trading in the thesis towel for

the graduate paper towel, they were able to knock some sense into me and keep me on track.

I would also like to thank my family who has always supported everything I've done. Certainly none of this would have been possible if I didn't have the mental and emotional support from my parents and siblings.

Lastly, I'd like to thank the Gleason Foundation for funding this research.

ABSTRACT

Popular semiconductors currently being used for RF applications include GaAs and InP. The operating frequencies for HEMTs built with these semiconductors covers the range from 800 MHz to 100 GHz. Although high-speed operation is attainable using GaAs or InP, product performance is limited when considering high-power applications, where a high breakdown field and thermal conductivity is needed. Proposed about fifteen years ago as a solution to this problem, GaN has emerged as a viable candidate to challenge, and perhaps overtake, GaAs and InP as the dominant semiconductor in RF products, particularly power amplifiers. This is due to its high breakdown field, large band gap, and high thermal conductivity (relative to InP and GaAs). GaN possesses other material parameter values that are favorable to existing technologies such as carrier saturation velocity, dielectric constant, and piezoelectric coefficients. On a worldwide scale, researchers and industry experts continue to work on the modeling of GaN-based HEMT devices in hope that GaN can be used commercially in the near future.

The proposed model is physics-based, making use of the Schrödinger and Poisson equations to establish relationships between the sheet carrier density, Fermi Level, and device terminal voltages. The quantum well formed at the heterointerface is approximated as a triangular well with two eigenstates, both determined from the Schrödinger equation. The charge control equations were carried over into the I/V derivations and channel charge derivations for capacitance calculations.

Spontaneous and piezoelectric polarizations at the heterointerface are responsible for the high density of carriers in the channel and are accounted for in the model in the expression for the device threshold voltage. Device performance is dictated by the aluminum mole fraction of the barrier layer. This is because the mole fraction controls the amount of polarization at the heterointerface and consequently the 2DEG density.

I/V equations were derived incorporating both drift and diffusion components. A two-region model was adopted for the saturation region to account for channel length modulation. Device conductances were derived from the drain current expressions and results compared to experimental data gathered.

To address high frequency device behavior, parasitic gate capacitance (C_{gs} and C_{gd}) expressions and cutoff frequency expressions are derived and presented. High voltage conditions were assumed for the drain bias to simulate a high power scenario. Relationships between the cutoff frequency of the device, the length of the gate, and drain bias are shown and compared with published data reported by other authors.

Table of Contents

List of Tables

List of Acronyms

Acronym	Meaning
2DEG	2-Demensional Electron Gas
Al	Aluminum
AlGaN	Aluminum Gallium Nitride
BJT	Bipolar Junction Transistor
CW	Continuous Wave
FET	Field Effect Transistor
GaAs	Gallium Arsenide
GaN	Gallium Nitride
HBT	Heterojunction Bipolar Transistor
HEMT	High Electron Mobility Transistor
HFET	Heterojunction FET
ID	Intentionally Doped
InP	Indium Phosphide
LHS	Left Hand Side
MESFET	Metal Semiconductor FET
MODFET	Modulation Doped FET
MOS	Metal Oxide Semiconductor
MOSFET	Metal Oxide Semiconductor FET
MS	Metal Semiconductor
RHS	Right Hand Side
SDFET	Selectively doped FET
Si	Silicon
SiC	Silicon Carbide
SiO_2	Silicon Dioxide
TEGFET	Two Dimensional Electron Gas FET
UID	Un-Intentionally Doped
VGA	Video Graphics Array

List of Symbols

Symbol	Meaning	Symbol	Meaning
$\varepsilon(m)$	AlGaN Dielectric constant	k_B	Boltzmann's Constant
ε^*	Basal Strain at heterointerface	L	Gate Length
ε_0	Permittivity Constant	L_1	Length of Low-Field Region
ε_{AlN}	Dielectric Constant of AlN	L_2	Length of High-Field Region
ε_{GaN}	Dielectric Constant of GaN.	L_{DS}	Drain to Source Contact Length
$\mu(x)$	Field Dependent Mobility	L_{GD}	Gate to Drain Contact Length
μ_0	Low-Field Mobility	m	Aluminum Mole Fraction
π	Pi	m^*	GaN Electron Effective Mass
$\sigma(m)$	Polarization-induced Sheet Charge Density	m_0	Electron Rest Mass
$\varphi_m(m)$	Schottky Barrier Height	$m_{e,AlGaN}$	AlGaN Electron rest Mass
$a_{AlGaN}(m)$	AlGaN Lattice Constant	$m_{e,AlN}$	AlN Electron Rest Mass
a_{GaN}	GaN Lattice Constant	N_D	Barrier Layer Doping Density
a_{AlN}	AlN Lattice Constant	n_s	Sheet Carrier Density
C_{gs}	Gate-to-Source Capacitance	n_{s0}	Equilibrium Sheet Carrier Density
C_{gd}	Gate-to-Drain Capacitance	n_{sat}	Saturation Sheet Carrier Density
Δd	Effective Width of 2DEG	P_D	Power Density
d_{cap}	Thickness of Cap Layer	$P_{piezo}(AlGaN)$	AlGaN Piezoelectric Polarization
d_d	Thickness of Doped Barrier Layer	$P_{spont}(AlGaN)$	AlGaN Spontaneous Polarization
d_i	Thickness of Spacer Layer	$P_{spont}(GaN)$	GaN Spontaneous Polarization
ΔE_c	Conduction Band Discontinuity	q	Electron Charge
$E(x)$	Position Dependent Electric Field	Q	Channel Charge
E_0	Lowest Eigenstate in Channel	R_0	Output Resistance
E_1	Second Lowest Eigenstate in Channel	R_s	Parasitic Source Resistance
E_c	Critical Electric Field	R_d	Parasitic Drain Resistance
E_c	Conduction Band (In diagram only)	T	Temperature
E_f	Fermi Level	V	Volts
E_v	Valence Band	$v(x)$	Channel Carrier Velocity
$E_{g,AlGaN}$	AlGaN Energy Gap	$V_c(x)$	Channel Potential
$E_{g,AlN}$	AlN Energy Gap	V_{ds}	Drain-to-Source Bias
$E_{g,GaN}$	GaN Energy Gap	V_{gs}	Gate-to-Source Bias
f_T	Cutoff Frequency	V_H	Voltage across High-Field Region
f_{max}	Max. Frequency of Oscillation	v_{sat}	Electron Saturation Velocity
g_m	Transconductance	V_T	Thermal Voltage
g_0	Output Conductance	$V_{th}(m)$	Threshold Voltage
h	Planck's Constant	W	Gate Width
h_{21}	Forward Gain	x	Channel Location
I	Current	x'	Modified Channel Location
$I_{d,lin}$	Linear Region Drain Current	y	AlGaN barrier depth location
I_{ds}	Drain Current		
I_{dsat}	Saturation Region Drain Current		

PUBLICATIONS

1. Jonathan C. Sippel, Syed S. Islam and S. S. Mukherjee, "A physics-based analytical model of a GaN/AlGaN HEMT incorporating spontaneous and piezoelectric polarization," Accepted for presentation in *IEEE Canadian Conference on Electrical and Computer Engineering*, 2004.

2. Jonathan C. Sippel and Syed S. Islam, "A charge control model for AlGaN/GaN HEMTs incorporating spontaneous and piezoelectric polarizations," Proceedings 27th Annual EDS/CAS Activities in Western New York Conference, p. 18, November 2003.

3. Jonathan C. Sippel and Syed S. Islam, "A physics-based model for AlGaN/GaN HEMTs," To be submitted.

Chapter 1

Introduction and Historical Review

1.1 Introduction

The wireless telecommunications industry has exploded over the past decade. This is due to the increased demand for low-cost high-volume products to be developed at a rapid pace. The rapid development of such technology has led to smaller circuit sizes, lower power supply ranges, and increased product features. The ever-expanding demands of consumers have led companies and researchers to find semiconductor technologies capable of meeting such demands. Contemporary wireless semiconductor technologies based on Silicon (Si), Gallium Arsenide (GaAs), and Indium Phosphide (InP) are used to implement various products such as video graphics array (VGA) camera sensor modules, driver-stage amplifiers, and integrated transceivers. These semiconductors have been used to design circuits with remarkable performance characteristics with wide-ranging functionality. Each, however, has been shown to be limited in its ability to handle high frequency, high power applications operating at temperature extremes.

Contemporary devices that have been implemented using the aforementioned semiconductors include Metal Oxide Semiconductor Field Effect Transistors (MOSFETs), Metal Semiconductor Field Effect Transistors (MESFETs), Bipolar Junction Transistors (BJTs), Heterojunction Bipolar Transistors (HBTs), and High Electron Mobility Transistors (HEMTs). These devices have proven their capability to implement many different types of products, though their field of application is limited to those products which do not require simultaneous high-power, high-frequency, high-temperature operation. The following sections will introduce a semiconductor and device technology capable of operating in all three areas (high temperature, power, frequency) simultaneously. The proposed semiconductor is Gallium Nitride (GaN) and the device is HEMT.

The following variables all refer to DC quantities and not AC quantities as their subscripts would imply otherwise: V_{gs}, V_{th}, V_{ds}, I_{ds}, $I_{d,lin}$, $I_{d,sat}$, and $V_c(x)$.

1.2 HEMT: What is it?

HEMT is an acronym for High Electron Mobility Transistor. Like other devices, the HEMT has numerous "aliases" such as MODFET (Modulation Doped Field Effect Transistor), SDFET (Selectively Doped Field Effect Transistor), HFET (Heterojunction Field Effect Transistor), and TEGFET (Two-Dimensional Electron Gas Field Effect Transistor) [1]. Each alias describes either the structure of the transistor or some characteristic of it. For example, from the aliases provided it can be gathered that the structure is not entirely made from one semiconductor and that current conduction is carried out through a two-dimensional electron gas (2DEG). Figure 1.1 shows the epitaxial structure of a HEMT transistor in its entirety, complete with source and drain terminals.

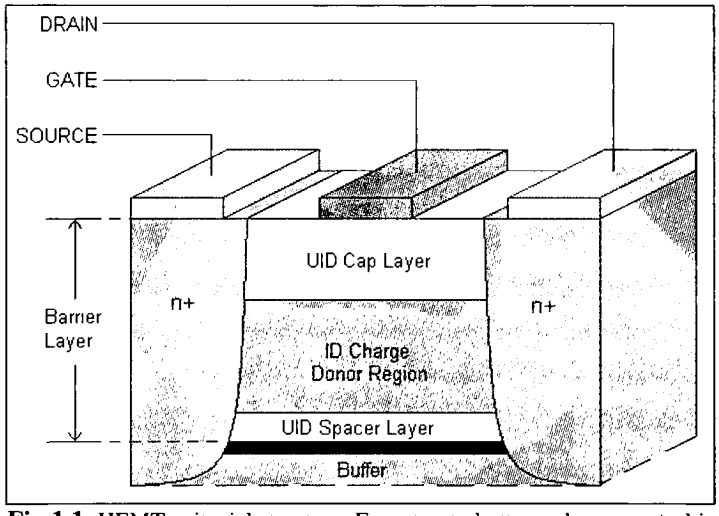

Fig 1.1. HEMT epitaxial structure. From top to bottom, charge control is made possible via a Schottky contact, followed by a cap layer, charge donor region, spacer layer, and buffer. The buffer is grown on top of a nucleation (substrate) layer (not shown in the picture).

At first glance, the HEMT structure pictured in Fig. 1.1 is reminiscent of other FET devices. In actuality, the HEMT is a combination of the MOSFET and MESFET. In a MOSFET structure, the gate is separated from the channel by an insulator, allowing for charge control over the channel. In a HEMT, the gate is separated from the channel by a barrier layer. The barrier layer is made of an alloy of the buffer

semiconductor and is shown in the figure to have three sub layers: 1) Un-intentionally doped (UID) Cap Layer, 2) Intentionally Doped (ID) Charge Donor Layer, and 3) UID Spacer Layer. The barrier layer and its sub layers will expounded upon later. The current channel is located in the buffer layer. One of the differences between the MOSFET and HEMT structures lies in the fact that the charge control in a MOSFET occurs via a MOS capacitor whereas in a HEMT the charge control is through a Schottky contact. The big difference is that a MOS capacitor can be used over a wide range of gate voltages, limited only by the breakdown voltage of the insulator between the gate and channel. This is not the case with the Schottky contact whose gate voltage range is limited by the barrier height of the metal. The height of the barrier depends on the type of metal the gate is made of and the semiconductor the gate is placed in contact with. Another difference between the MOSFET and HEMT devices concerns the carrier confinement in the channel. Assuming a MOSFET fabricated in Si, the energy difference between the SiO_2 insulator and Si substrate is on the order of a few electron volts (eV) resulting in exemplary carrier confinement. The HEMT device does not use an insulator and relies on heterointerface polarization fields and interface conduction band discontinuity for carrier confinement. The band discontinuity is less than 1 eV, resulting in a worsened carrier confinement relative to Si. For the same reason a HEMT is different from a MOSFET, the HEMT is similar to a MESFET. Both the MESFET and the HEMT employ a Schottky contact to mediate current flow in the device. The difference between these devices is how the mediation is performed. For example, in a MESFET there is no barrier or insulation layer between the gate and channel so current flows directly beneath the gate between the source and drain terminals. The magnitude of current is regulated by the amount of reverse bias applied between the gate and source terminals modulating the depletion region width beneath the gate. The depletion region extends into the current path controlling the volume of current flowing through the device. From a device standpoint, this is fundamentally different than a HEMT which operates more like the MOSFET.

There is another difference between the HEMT and MESFET structures other than how the current is mediated in the devices. In a MESFET, a Schottky contact is placed on top of a doped semiconductor channel and is reverse biased to control the amount of current flowing in the device. The channel is located directly beneath the gate in the doped semiconductor. So, if one wishes to design a structure capable of

higher current density, one would be inclined to increase the doping density of the semiconductor. The problem with that is that, as the doping density increases, the mobility of the current-carrying electrons decrease as more and more carriers are squeezed into a static-sized region. So although the overall goal of increasing current density may be realized, the magnitude of change will not be as pronounced as if the mobility of the carriers were able to maintain its value. This is the problem the HEMT was proposed to fix with its introduction by Dr. Takashi Mimura in the late 1970's [2]. By employing a modulation doping scheme, the HEMT physically separates the current channel from the doped semiconductor, allowing for the carrier density in the channel to benefit from a doped semiconductor region without suffering from reduced carrier mobility due to impurity scattering.

A discussion of the band diagram of the device may help to clarify some of the device aspects mentioned thus far. Fig. 1.2 depicts the band diagram corresponding to the structure shown in Fig. 1.1.

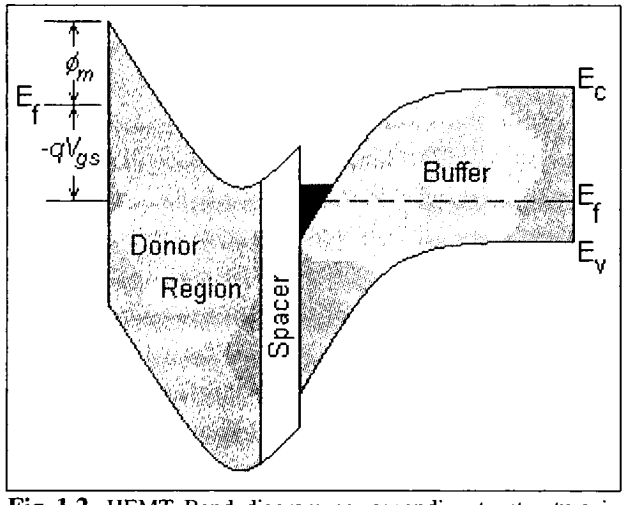

Fig 1.2. HEMT Band diagram corresponding to structure in Fig. 1.1.

The band diagram has been appropriately labeled to show the relationship between the figures. Progressing from left to right, the diagram shows how the energy bands of the semiconductor bend to account for internal electric fields, band discontinuity, and depletion regions created at the different interfaces. The metal-semiconductor (MS) contact at the left side of the figure results in a depletion region on the semiconductor (donor region) side of the junction and an associated barrier height, φ_m. Since the donor region is doped n-type, the conduction band, E_c, is observed to slope downward as the distance from the

MS interface increases. The spacer layer is undoped so electrons from the donor region diffuse into the spacer layer causing the conduction band to bend upwards. Continuing rightward, an interesting phenomenon is observed where the spacer layer meets the buffer layer. Earlier it was stated that the buffer and barrier layers had the same base semiconductor type, but that the material making up the barrier layer had a defined alloy composition. This alloy composition causes the energy gap of the barrier layer to be larger than that of the buffer layer resulting in a conduction band discontinuity, ΔE_c, at the interface. This discontinuity results in the formulation of a triangular potential well directly adjacent to the interface. In this potential well, a two-dimensional electron gas (2DEG) forms (indicated by the dark buildup in the well beneath the Fermi Level); "two-dimensional" refers to the reduced dimensionality of the electron momentum inside the accumulated region [1]. The density of the 2DEG depends on two components: 1) Donor component, 2) Polarization component. The donor component is directly related to the thickness and doping density of the charge donor region in the barrier layer. When the device is fabricated an equilibrium state is created in the absence of any external bias. The Fermi level aligns between all regions and all mobile carriers are depleted from the barrier layer directly below the gate. The depleted carriers empty into the potential well and become trapped by potential barriers on either side. At this time it is pertinent to add that the barrier layer must be thin enough so that the depletion regions of the MS interface and heterointerface overlap so the semiconductor area below the gate is entirely depleted. If they do not overlap and mobile carriers are left in the region, a conducting channel may form in the barrier layer in the presence of an applied drain to source bias, V_{ds}. This parasitic MESFET channel will degrade the performance of the device to such an extent that it will be rendered unusable. The other component adding to the 2DEG density is the polarization component (assuming the buffer layer semiconductor to be piezoelectric). The polarization component is made up of two sub-components: 1) Piezoelectric component, and 2) Spontaneous component. The piezoelectric component is created from the tensile strain resulting from the differences in lattice constants between the buffer and barrier layers. The tensile strain creates an electric field pulling more charges into the channel and raises the concentration of the 2DEG. Increasing the alloy composition in the barrier layer increases the amount of tensile strain at the interface and, consequently, the number of induced charges. It is important to mention here that the alloy composition, if too high, can

cause bonds to break at the interface, forming misfit dislocations resulting in performance degradation. If the alloy composition is so high that bonds break, the strain at the interface is said to be relaxed (less than full strain). The optimal scenario is to have full strain at the interface resulting in the highest amount of piezoelectric polarization-induced charges. The alternate polarization sub-component, termed spontaneous polarization, results from the cation and anion positions in the lattice [3]. Differences between the spontaneous polarizations of two adjacent layers can create a high density of mobile carriers. Both polarization effects create electric fields at the heterointerface which help to confine the carriers to the potential well. With the reasoning behind the potential well having been thoroughly discussed, attention can now return to the band diagram. As one progresses further right extending into the buffer layer, the band diagram is observed to flatten out and become constant as expected.

As will be demonstrated in this thesis, V_{gs} is ultimately responsible for the 2DEG concentration level in the potential well so discussing the relationship between V_{gs} and the Fermi Level, E_f, is pertinent to understanding device behavior presented in the theory section. Fig. 1.2 provides a visual representation of V_{gs}'s effect on the system. It is important to note here that any change in V_{gs} does not change the band bending in the barrier layer. Any change in V_{gs} shifts the energy bands of the entire barrier layer either up, if V_{gs} is negative, or down, if V_{gs} is positive, thereby modulating the distance between the bottom of the conduction band discontinuity and the Fermi Level, E_f. It is this distance between the bottom of the conduction band and E_f which dictates the magnitude of the 2DEG concentration. In the next chapter, relationships will be presented between the 2DEG sheet carrier density, n_s, and E_f and n_s and V_{gs}.

1.3 Why GaN?

In recent years Gallium Nitride (GaN) has emerged as a viable candidate to replace the "wireless" semiconductors listed in Section 1.1 whose performance capabilities limited them to products that did not require simultaneous high frequency, high power, high temperature operation. High Electron Mobility transistors on GaN substrates have demonstrated an ability to handle the operational modes collectively better than the aforementioned semiconductors. Table 1.1 lists the material parameters for GaN, SiC, Si, GaAs, and InP for comparison. It is clear from the table that GaN collectively is a very appealing

semiconductor technology in that it blends "the best of everything" – high frequency operation, high temperature capability, with high power capability. According to Table 1.1, GaN has the highest breakdown field and bandgap at 5 MV·cm^{-1} and 3.42 eV, respectively. These material parameters are a testament to its high power density, P_D, capability revealing its ability to handle severe biasing conditions. The mobility, dielectric constant, and carrier saturation velocity reveal the ability of GaN to operate at high frequencies. The relative dielectric constant is 10 and the saturation velocity is 1.45x10^7 cm/sec, which is second only to SiC. The electron mobility is superior to that of SiC. The low dielectric constant means that any parasitic capacitances in the device will have less of an effect on the intrinsic frequency response of the device. For instance, having a lower dielectric constant helps to lower the values of the gate capacitance, gate-to-source capacitance, C_{gs}, and gate-to-drain capacitance, C_{gd}. Lowering the aforementioned capacitance values raises the cutoff frequency, f_T, and maximum oscillation frequency, f_{max}, of the device. f_T and f_{max} are figures of merit used to define the switching speed and power gain of a device, respectively [1].

Table 1.1. Semiconductor material parameter comparison [4,5].

Parameter	GaN	SiC	Si	GaAs	InP
Breakdown Field (V·cm^{-1})	5x10^6	1x10^6	3x10^5	4x10^5	5x10^5
Bandgap (eV)	3.42	2.36	1.11	1.43	1.344
Electron Mobility (cm^2/V·sec)	1100	900	1400	8500	5400
Relative Dielectric Constant	10	9.66	11.7	12.5	12.5
v_{sat} (x10^7 cm/sec)	1.45	2.5	1	0.7	1.5
Thermal Conductivity (W/cm·K)	1.7	4.9	1.3	0.54	0.68

The thermal conductivity of 1.7 W/cm·K is second only to SiC, allowing GaN to maintain performance levels at elevated temperatures. GaN HEMT devices have even been grown on SiC substrates to benefit from the higher thermal conductivity, therefore enhancing performance.

Another aspect of GaN which makes it very attractive is that it is a piezoelectric material. It, therefore, benefits from the polarization effects mentioned in the previous section. These effects help to raise the concentration of the 2DEG a great deal. When compared to GaAs-based HEMT structures, the 2DEG concentration has been observed to be more than five times larger [6] leading to increased current and power densities. The most commonly used alloy in the barrier layer to create the lattice strain at the heterointerface is Aluminum (Al). Most authors have referred to such structures as AlGaN/GaN HEMTs, or simply AlGaN HEMTs indicating what the barrier layer is made of. Authors will sometimes opt not to

explicitly say what the aluminum composition is in the barrier layer, instead indicating it as a subscript to the acronym for the device as: $Al_mGa_{1-m}N$, where m is called the aluminum mole fraction and is between 0 and 1 with a 0 corresponding to pure GaN and 1 corresponding to pure AlN.

References [7,8,9] report physical samples that have been fabricated and tested for device performance. Testing of these devices revealed continuous wave (CW) power density values, P_D, as high as 16.5 W/mm with f_{max} and f_T values as high as 100.9 GHz and 50 GHz, respectively. In [8], a power HEMT was fabricated and shown to withstand biasing up to 600 V legitimizing itself as a power device for motor drive and power applications. In [7], drain biasing went as high as 60 V revealing the GaN-based HEMTs effectiveness in being considered for power amplifier applications. The results published by these authors, along with the steady improvement in growth techniques for GaN, has many researchers and companies eagerly awaiting its commercial availability.

1.4 Contemporary Modeling Issues

The growth and modeling issues associated with GaN and GaN-based HEMTs are undoubtedly the reason which has thus far kept it from wide-spread use commercially. Steady progress has been made each year, although more is needed until mass production of GaN products can begin. Contemporary issues under scrutiny include:

1) Modeling of thermal and trapping effects.
2) Correct incorporation of polarization term in device threshold voltage equation.
3) Modeling of growth process, and
4) I/V current modeling incorporating both diffusion and drift current components in linear and saturation regions of operation.

The above issues are of utmost importance if GaN-based HEMTs are to be used to develop high-speed wireless products. The "thermal effect" refers to the mobility degradation of the carriers as the device heats up. This effect is observed as a drop in drain current with increasing drain bias. Correct modeling of the phenomenon is essential for predicting device behavior under high bias, high temperature conditions. The "trapping effect" refers to electron traps incorporated in the structure during material growth and device

fabrication. The trapping effects include transconductance frequency dispersion, current collapse, light sensitivity, and gate- and drain-lag transients [10]. Understanding these trapping effects and how to correctly model them allows for accurate device simulations to be performed and increases the reliability of circuit simulations that will be performed when GaN-based HEMTs are incorporated into circuit designs. Although encouraging work has been performed with regards to modeling the polarization induced charge [6,11], more work is needed to model how the polarization effects at the heterointerface affect the threshold voltage of the transistor for elevated mole fractions. Many I/V models have been reported to predict current levels in both the linear and saturation regions of operation but none, to the best of the author's knowledge, have included both drift and diffusion components in the derivation of the linear and saturation region current equations. Such a model would be valuable in that it could be used to derive small-signal device parameters and provide insight into device behavior under different biasing conditions.

The modeling issues outlined in this chapter provide the reasoning why GaN has not seen wide commercial use. GaN usage is sure to increase in the near future as these issues become well understood by researchers.

1.5 Literature Review

Several existing models, equations, and constants were used in this thesis to produce the results presented later. All of the borrowed literature was chosen because of its proven accuracy. This does not mean, however, that all the adopted models and/or equations are 100% accurate. This chapter details what models have been adopted in this thesis, why they were adopted, and what their limitations are. This is done to give the reader an appreciation for why certain parameters or expressions were used and to provide motives for future research.

In [1,12] a relationship is derived between the sheet carrier density, n_s, and the Fermi Level, E_f. The flow of the derivations differed slightly from one another, although the end result is the same. Both derivations assumed the quantum well at the heterointerface to be triangular and mentioned the use of the Schrödinger equation to generate the quantum well eigenstates. The final equation detailing the relationship between n_s and E_f was found, in both cases, to be [1,12],

$$n_s = Dk_BT \ln\left[1+e^{\frac{E_f-E_0}{k_BT}}\right] + Dk_BT \ln\left[1+e^{\frac{E_f-E_1}{k_BT}}\right] \quad (1.1)$$

where k_B is Boltzmann's constant, T is temperature in Kelvin, E_0 and E_1 are the quantum well eigenstates, and D is a constant equal to $4\pi m^*/h^2$ (h is Planck's constant and m^* is the electron effective mass in GaN). With only two eigenstates considered, equation (1.1) is relatively easy to formulate and does produce reasonably accurate results. However this strength is also the equation's greatest weakness. Since the number of eigenstates has been limited to two, the accuracy of it is limited to certain regions. Most don't consider this "weakness" when the device is biased in the strong inversion region since the lower eigenstates are highly populated.

References [1,13,14] have provided linear charge-control models relating n_s to V_{gs} using the Poisson equation. The relationship is shown in equation (1.2),

$$n_s(m) = \frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)}\left(V_{gs} - V_{th}(m) - \frac{E_F}{q}\right) \quad (1.2)$$

where $\varepsilon(m)$ is the dielectric constant of the barrier layer, q is the electron energy, d_d is the thickness of the charge donor region in the barrier layer, d_i is the thickness of the undoped spacer region in the barrier layer, Δd is the effective width of the 2DEG, V_{gs} is the applied gate bias, and V_{th} is the threshold voltage. When biased in the strong inversion region, equation (1.2) has been shown to be quite accurate in its n_s concentration predictions [15]. Since (1.2) is a linear equation, its accuracy decreases for V_{gs} values near threshold and below where n_s has been experimentally observed to decrease at an exponential rate [16]. This weakness was deemed acceptable since the transistor was biased to operate in the strong inversion region only.

The I/V model that was adopted closely followed that reported by Rashmi *et al.* [17,18]. It was decided that this model was the best to start since it incorporated drift and diffusion components in the linear region [17], accounted for channel length modulation in the saturation region by using the 2D Poisson equation [18], and incorporated a reliable carrier mobility model reported by Ruden *et al.* [19]. The problem with the current model reported in [17] is that it does not account for channel length modulation in the saturation

region, assuming drain current to remain constant with increasing drain bias. The problem with the model reported in [18] is that it neglects to incorporate the diffusion current component altogether. Although these papers reported good results, a revised model should be made that accounts for both drift and diffusion current components, and channel length modulation.

The threshold voltage model that was used was proposed by Rashmi *et al.* [15]. In [15], V_{th} is given to be dependent on the magnitudes of the polarization fields induced at the heterointerface as,

$$V_{th}(m) = \varphi_m(m) - \Delta E_c - \frac{qN_D d_d^2}{2\varepsilon(m)} - \frac{\sigma(m)}{\varepsilon(m)}(d_d + d_i) \tag{1.3}$$

where φ_m is the Schottky barrier height, ΔE_c is the conduction band discontinuity at the hetero-interface, N_D is the doping density, and $\sigma(m)$ is the polarization-induced sheet charge density. Using (1.3), V_{th} can theoretically be calculated for a device of specified geometry and mole fraction. The author found it difficult to calculate reasonable V_{th} values using (1.3) for devices of mole fractions in excess of 0.15, limiting the number of experimental samples that could be used for data comparison. The model that was adopted for $\sigma(m)$ was reported by Ambacher *et al.* [6] and showed excellent agreement between calculated and experimental results, substantiating the polarization model's credibility. Attempts were made to locate other authors who may have tried to obtain a relationship between V_{th} and $\sigma(m)$ but very little has been done up to this point. The inability of (1.3) to reliably calculate V_{th} values for devices with high aluminum content was found to be a huge setback in that it limits the scope of the model to a narrow range of devices. Further research is needed to fully understand the polarization effect on the device threshold voltage.

The models just discussed were used in this thesis for one of two reasons: 1) They provide accurate predictions for the biasing conditions considered, or 2) They can be improved upon. In either case, they provided an excellent starting point for understanding device behavior along with what past researchers have done to model that behavior.

1.6 Contributions

The contributions of this thesis are as follows:

1) A relationship between the sheet carrier density, n_s, and Fermi level, E_f, incorporating two quantum well eigenstates is shown. Plots are given to visually show n_s's dependence on E_f.
2) A relationship between the applied gate voltage, V_{gs}, and the sheet carrier density is presented. The relationship was derived from the Poisson equation and aims to provide an understanding of how V_{gs} affects n_s.
3) Relationships between n_s and the barrier thickness and V_{th} and the doping density are presented with discussion to provide understanding of how changing aspects of the barrier layer will change device performance.
4) Linear and saturation region drain current equations are derived. Drift and diffusion components are included in both regions. Channel length modulation has also been incorporated in the saturation region. Predicted results are compared with experimentally measured data.
5) Device transconductance and output conductance have been derived. Both are plotted for various bias conditions whose plots are coupled with discussion to provide insight into the results calculated by the derived model. Conductances were derived using the current equations so both drift and diffusion components are incorporated.
6) Expressions for gate-to-source and gate-to-drain capacitance are derived. When the device was biased in the saturation region, the channel charge was calculated using a two-region model (High-Field region and Low-Field region). The device capacitances are plotted against applicable bias conditions. Discussion is provided to interpret the results and provide insight into the device's high power, high frequency capabilities.
7) An expression for the intrinsic cutoff frequency of the device is presented and relationships are established between the cutoff frequency of the transistor, length of the gate, and the drain bias.

1.7 Thesis Organization

This thesis has been partitioned into five chapters: 1) Introduction and Historical Review, 2) Theory / Model Formulation, 3) Materials and Apparatus, 4) Results and Discussion, and 5) Conclusions and Future Work. Each of these chapters provides information pertinent to the understanding of the proposed material

along with what future modifications can be made to extend the work further. Subsections are included in the larger chapters, such as this one, where many important topics are presented and discussed. The chapters just mentioned are listed below with a brief description detailing their contents.

Chapter 1 – "Introduction and Historical Review" provides all background information needed to form a basic understanding of what a HEMT device is, how it operates, and how it compares to other structures, notably a MOSFET and MESFET. Background is given about the GaN semiconductor and why it is favorable when compared with other leading wireless semiconductors. The motivation for the thesis along with what contributions have been given to the electrical engineering field is also discussed. Chapter 2 – "Theory / Model Formulation" details the derivational process for most equations used. Some equations that were borrowed from other authors were not re-derived since the sources can be referenced directly. Other equations that had lengthy derivations listed only the final equation with the intermediate steps of the derivation saved for the appendix attached at the end. Chapter 3 – "Materials and Apparatus" is a semi-formal discussion detailing all materials used to complete this thesis. The reference used for experimental data is also mentioned along with the dimensions of the device that were reported in the paper. Chapter 4 – "Results and Discussion" presents results for each section in Chapter 2. A complete analysis is provided for each figure and additional insight is offered to reinforce important points. In Chapter 5 – "Conclusions and Future Work" the results of the thesis are wrapped up into a brief summary highlighting the strengths and weaknesses of the proposed model. Major weaknesses are again mentioned to provide a foundation for future research in this area.

Chapter 2

Theory / Model Formulation

2.1 Introduction

This chapter presents two models that are capable of predicting 2DEG carrier concentration levels. One of the 2DEG models is later used to derive drain current equations and associated small-signal parameters such as transconductance and parasitic gate capacitances. Of the two models presented, the first relates the 2DEG concentration to the Fermi level using Fermi-Dirac statistics and the second relates the 2DEG concentration to the applied gate-to-source bias using the Poisson equation. The high electron mobility transistor *I/V* characteristic curve equations are derived next followed by a detailed derivation of the small-signal parameters.

2.2 Sheet Carrier Density, n_s, vs. Fermi Level, E_f

The performance of the HEMT device is centered on the concentration of the 2DEG. Therefore a model must be developed that can accurately and reliably predict the 2DEG concentration. Multiplying the density-of-state by the probability of occupancy function for finding an electron at a particular energy level E, an equation for the channel electron density can be [1],

$$n_s = \int_0^\infty D(E) f(E) dE \tag{2.1}$$

where $D(E)$ is the density-of-state function and $f(E)$ is the probability function for finding an electron at a particular energy E. Employing Fermi-Dirac statistics, $f(E)$ becomes [1],

$$f(E)=\frac{1}{1+e^{\frac{E_f-E_c}{k_BT}}} \tag{2.2}$$

where k_B is Boltzmann's constant, T is the temperature in Kelvin, E_f is the Fermi level, and E_c is the conduction band energy level. Fermi-Dirac statistics account for two restrictions: (1) It is impossible to distinguish one electron from another, and (2) No two electrons may occupy the same energy state, referred to as the Pauli exclusion principle [1]. Since the minimum energy valley in GaN, the Γ valley, is spherical with respect to momentum, the valley can be characterized by a single effective mass, m^*. Therefore the density-of-state function can be defined as follows (considering only the two lowest energy levels in the quantum well, E_0 and E_1),

$$D(E)=\begin{cases}0 & E\le E_0\\ \frac{4\pi}{h^2}m^* & E_0\le E\le E_1\\ 2\cdot\frac{4\pi}{h^2}m^* & E_1\le E\end{cases} \tag{2.3}$$

where E_0 is the lowest subband energy level, E_1 is the second-lowest subband energy level, h is Planck's constant, and m^* is the electron effective mass in GaN. Plugging (2.2) and (2.3) into (2.1) yields [1],

$$n_s=\frac{4\pi}{h^2}m^*\int_{E_0}^{E_1}\frac{1}{1+e^{\frac{E-E_f}{k_BT}}}dE+\frac{8\pi}{h^2}m^*\int_{E_1}^{\infty}\frac{1}{1+e^{\frac{E-E_f}{k_BT}}}dE$$

Performing the integration yields an equation relating the channel carrier density, n_s, to the Fermi level, E_f,

$$n_s=Dk_BT\ln\left[1+e^{\frac{E_f-E_0}{k_BT}}\right]+Dk_BT\ln\left[1+e^{\frac{E_f-E_1}{k_BT}}\right] \tag{2.4}$$

where $D = 4\pi m^*/h^2$ Using Schrödinger's equation and assuming a triangular potential well, expressions for the first two subband energy levels, E_0 and E_1, were found by Rashmi *et al.* [15] to be,

$$E_0=2.123\times10^{-12}\left(n_s^{2/3}\right)$$

$$E_1=3.734\times10^{-12}\left(n_s^{2/3}\right)$$

Fig. 2.1 shows a close up of the triangular quantum well at the heterointerface. The incorporated eigenenergies, E_0 and E_1, are shown in the quantum well along with the Fermi level.

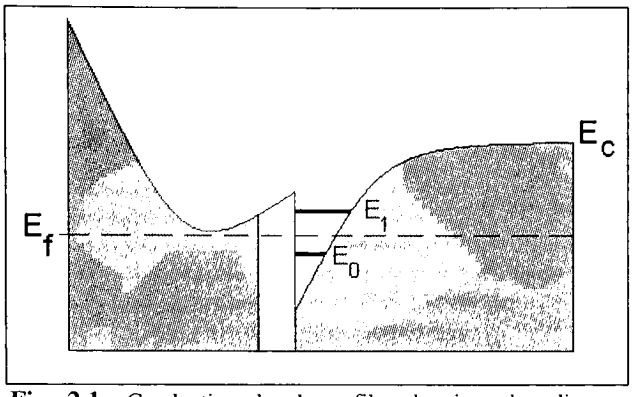

Fig 2.1. Conduction band profile showing the discrete eigenenergy levels E_0 and E_1. E_f shifts up and down with applied V_{gs} filling or emptying the energy levels with electrons.

If the Fermi level increases, eigenstates E_0 and E_1 can be expected to fill with electrons raising the 2DEG concentration. Conversely, E_0 and E_1 can be expected to be emptied of their electrons if the Fermi level drops.

Analysis of equation (2.4) reveals an exponential relationship between E_f and n_s for E_f values that cause the arguments of the exponential terms to evaluate to less than zero. In this range, E_f is well below both eigenenergy values of E_0 and E_1. Previous research indicates the range for n_s in this region is anywhere from 10^6 cm^{-2}, for very low values of E_f, to 10^{10} cm^{-2} for higher values of E_f. A linear dependence between n_s and E_f begins when E_f approaches E_0 in magnitude forcing the exponential term to evaluate to unity or larger. This region is known as the strong inversion region and is the target biasing region for the device for amplifier applications. n_s values in this region have been found to be on the order of 10^{11}~10^{13} cm^{-2} using the AlGaN/GaN heterostructure system.

To explicitly find the sheet carrier density as a function of the Fermi Level, equation (2.4) can be solved iteratively until convergence for a given Fermi level value. For example, if it was desired to find the equilibrium sheet carrier density, n_{s0}, E_f would be set equal to zero in equation (2.4) and an initial value would be used for n_s in the E_0 and E_1 expressions. The resulting E_0 and E_1 values would then be plugged into the right hand side (RHS) of equation (2.4) so that a numerical answer can be determined for the left hand side (LHS) of the equation. This new value for n_s is then re-substituted into the E_0 and E_1 expressions.

The new E_0 and E_1 values are again plugged into the RHS of equation (2.4) and the value for the LHS is solved for. This iterative process is repeated until the n_s, E_0, and E_1 values no longer change. When this happens, it can be said that the equation converged on a value for n_s for the given E_f value of 0 eV.

Having shown how equation (2.4) can be used to find an n_s value for a given E_f value, one can see how (2.4) can also be used to obtain a series of values for n_s for a specified range of E_f values. In the results section, a plot will be made detailing the relationship between n_s and E_f and insight will be given to shed some light on their relationship and how they are ultimately dependent on V_{gs}.

2.3 Sheet Carrier Density, n_s, vs. Gate Voltage, V_{gs}

Equation (2.4) shows the relationship between the sheet carrier density, n_s, and the Fermi level, E_f, and assumes a triangular potential well with only the first two quantum states (E_1 and E_0) occupied. Another equation for n_s can be obtained by using the Poisson equation in conjunction with the total depletion approximation. If it is assumed that the interface depletion region and gate depletion region overlap one another, and that total depletion occurs in the AlGaN barrier layer, simplifying and solving Poisson equation yields [15],

$$n_s(m) = \frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)}\left(V_{gs} - V_{th}(m) - \frac{E_f}{q}\right) \tag{2.5}$$

where m is the aluminum mole fraction, $\varepsilon(m)$ is the dielectric constant of AlGaN, d_d is the thickness of the doped AlGaN barrier layer, d_i is the thickness of the undoped AlGaN spacer layer, Δd is the effective width of the 2DEG, q is the electron charge, V_{gs} is the applied gate-to-source voltage, and $V_{th}(m)$ is the polarization-dependent threshold voltage given as [15, 17, 18],

$$V_{th}(m) = \varphi_m(m) - \Delta E_c - \frac{qN_D d_d^2}{2\varepsilon(m)} - \frac{\sigma(m)}{\varepsilon(m)}(d_d + d_i) \tag{2.6}$$

where φ_m is the Schottky barrier height, ΔE_c is the conduction band discontinuity at the hetero-interface, N_D is the doping density, and $\sigma(m)$ is the polarization-induced sheet charge density given as [20],

$$|\sigma(m)| = \left|P_{spont}(Al_mGa_{1-m}N) - P_{spont}(GaN) + P_{piezo}(Al_mGa_{1-m}N)\right| \tag{2.7}$$

where $P_{spont}(Al_mGa_{1-m}N)$ and $P_{spont}(GaN)$ are the spontaneous polarization of the AlGaN barrier layer and spontaneous polarization of the GaN buffer layer, respectively, and can be expressed as [20],

$$P_{spont}(Al_mGa_{1-m}N) = -0.052m - 0.029$$

$$P_{spont}(GaN) = -0.029$$

and $P_{piezo}(Al_mGa_{1-m}N)$ is the piezoelectric polarization of the AlGaN barrier layer and can be expressed as [21],

$$P_{piezo}(Al_mGa_{1-m}N) = mP_{AlN}^{pz}[\varepsilon(m)] + (1-m)P_{GaN}^{pz}[\varepsilon(m)]$$

where

$$P_{AlN}^{pz}[\varepsilon(m)] = -1.808\varepsilon^{*} - 7.888\varepsilon^{*2}$$

$$P_{GaN}^{pz}[\varepsilon(m)] = -0.918\varepsilon^{*} + 9.541\varepsilon^{*2}$$

ε^{*} is the basal strain and is related to the lattice constants of the GaN substrate and AlGaN barrier such that,

$$\varepsilon^{*}(m) = \frac{a_{GaN} - a_{AlGaN}(m)}{a_{AlGaN}(m)}$$

where the lattice constants, a_{GaN} and $a_{AlGaN}(m)$, were interpolated from experimental data using Vegard's Law as [21],

$$a_{Al_mGa_{1-m}N}(m) = 0.31986 - 0.00891m$$

The units of each polarization are in Coulombs per square meter (C/m^2). The piezoelectric polarization of the GaN buffer layer, $P_{piezo}(GaN)$, is assumed to be negligible since the buffer layer is much thicker than the barrier layer. If it were included, it would be subtracted from the AlGaN piezoelectric polarization term inside the absolute value brackets of equation (2.7). Equation (2.5) assumes no drain-to-source bias (V_{ds} = 0V) and models the total amount of charge depleted from the barrier layer.

The sheet carrier density can be determined as a function of V_{gs} according to equation (2.5). This was done by finding a best fit line to the linear portion of the n_s vs. E_f curve discussed in Section 2.2. Using linear regression analysis, a second-order best-fit line was obtained for the linear portion of the n_s vs. E_f curve as,

$$n_s = (7\times10^{14})E_f^2 + (9\times10^{13})E_f + 3\times10^{12}$$

Solving the above equation for E_f and plugging into equation (2.5) leaves an expression that can be solved for n_s as a function of gate voltage, V_{gs},

$$n_s(m) = \frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)}\left(V_{gs} - V_{th}(m) + \frac{9}{140} + \frac{1}{7\times10^7}\left(-7.5\times10^{11} + 7n_s\right)^{1/2}\right)$$

Solving for n_s,

$$n_s(m) = d\left(V_{gs} - V_{th} + a + b\left(\frac{7}{2}db - \frac{1}{2}\left(49d^2b^2 + 4c + 28dV_{gs} - 28dV_{th} + 28da\right)^{1/2}\right)\right) \quad (2.8)$$

where

$$a = \frac{9}{140} \quad \text{V}$$

$$b = \frac{1}{70000000} \quad \text{V·cm}$$

$$c = -7.5\times10^{11} \quad \text{1/cm}^2$$

$$d = \frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)} \quad \text{F/C·cm}^2$$

Equation (2.8) is the resulting linear charge control equation for n_s in terms of V_{gs}. It should be noted that since a best-fit line was used for the linear portion of the n_s vs. E_f curve, equation (2.8) will be accurate for predicting channel concentrations in the strong inversion region only, where a linear dependence is observed between n_s and V_{gs}. This will be discussed further in the results section.

2.4 Sheet Carrier Density, n_s, vs. Barrier Thickness, d_d

Equation (2.5) was used to show the relationship between sheet carrier density and barrier thickness. Both V_{gs} and V_{ds} were fixed at 0V so that the analysis could be performed under zero bias conditions. That left the threshold voltage, V_{th}, as the last unknown left in equation (2.5) which was readily calculated using GaN material parameters and device geometries provided in the subsequent chapter. E_f was assumed to be zero since its magnitude is negligible compared to the other terms in equation (2.6). Using the aforementioned assumptions and simplifications, n_s was calculated over a specified range for d_d.

2.5 Threshold Voltage, V_{th}, vs. Doping Density, N_D

Equation (2.6) was used to plot the relationship between the device threshold voltage, V_{th}, and barrier doping density, N_D. The same assumptions were used here as in Section 2.4, which is to say that V_{ds} and V_{gs} were set equal to 0V, and E_f to 0eV.

2.6 Current-Voltage Characteristics

The HEMT channel current can be modeled using the current density equation [15,17],

$$I_{ds} = Wq\mu(x)\left(n_s(m,x)\frac{dV_c(x)}{dx} + \frac{k_B T}{q}\frac{dn_s(m,x)}{dx} \right) \tag{2.9}$$

where x is the location in the channel with origin at the source side (refer to Fig. 2.2), W is the gate width, $V_c(x)$ is the channel potential, and $\mu(x)$ is the field-dependent mobility given in [19] as,

$$\mu(x) = \frac{\mu_0}{\left(1 + \frac{1}{E_1}\frac{dV_c(x)}{dx}\right)} \tag{2.10}$$

with

$$E_1 = \frac{E_c v_{sat}}{\mu_0 E_c - v_{sat}}$$

where E_c is the critical electric field, v_{sat} is the saturation drift velocity of electrons, and μ_0 is the low-field mobility. The first term in the parentheses of equation (2.9) accounts for the drift component and the second term for the diffusion component of the current. Neither of the two previous charge control equations, (2.8) or (2.5), can be used in equation (2.9) for n_s because neither considers the charge density under drain bias conditions. The charge control model to be used here has been reported in [1,15,13] as,

$$n_s(m,x) = \frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)}\left(V_{gs} - V_{th}(m) - V_c(x)\right) \tag{2.11}$$

which includes the channel potential term and can be derived in the same manner as equation (2.5). Plugging (2.10) into (2.9) and cross-multiplying yields,

$$I_{ds}\left(1+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)\frac{dV_c(x)}{dx}\right) = Wq\mu_0\left(n_s(m,x)\frac{dV_c(x)}{dx} + \frac{k_B T}{q}\frac{dn_s(m,x)}{dx}\right) \tag{2.12}$$

The boundary conditions for equation (2.12) can be formed by knowing the intrinsic voltages at either end of the channel directly beneath the gate. In this case, parasitic source and drain resistances have been taken into account making the boundary conditions as follows:

$$V_c(x)\big|_{x=0} = I_{ds}R_s$$

$$V_c(x)\big|_{x=L} = V_{ds} - I_{ds}R_d$$

where R_s and R_d are the parasitic source and drain resistances respectively. Integrating equation (2.12) with respect to the $V_c(x)$ from source to drain and solving for I_{ds} yields

$$I_{d,lin} = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} \tag{2.13}$$

where

$$\alpha = (R_d + R_s)\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right) - \frac{W\mu_0\varepsilon(m)}{2d}\left(R_d^2 - R_s^2\right)$$

$$\beta = \frac{W\mu_0\varepsilon(m)}{d}V_{ds}R_d - \frac{W\mu_0\varepsilon(m)}{d}(R_D + R_s)V_{gs}' - L - \left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)V_{ds}$$

$$\gamma = \frac{W\mu_0\varepsilon(m)}{d}V_{gs}'V_{ds} - \frac{W\mu_0\varepsilon(m)}{2d}V_{ds}^2$$

$$d = d_d + d_i + \Delta d$$

$$V_{gs}' = V_{gs} - V_{th}(m) - \frac{k_B T}{q}$$

The intermediate steps of the derivation of equation (2.13) are shown in Appendix I. Equation (2.13) gives the current model for the linear region of operation. The condition that was used to determine linear region operation was,

$$\frac{\partial I_{d,lin}}{\partial V_{ds}} > \frac{\partial I_{dsat}}{\partial V_{ds}}, \quad V_{gs} > V_{th}$$

where I_{dsat} refers to the saturation drain current to be derived next.

To get the saturation region drain current equation a different approach was used. In saturation, the channel is separated into two regions: (1) A Low-Field Region, and (2) A High-Field Region [18]. The high-field region occurs close to the drain side where the channel experiences pinch-off as shown in Fig. 2.2. The majority of the applied drain voltage will be dropped across this pinched off region and all current can be considered drift current due to the electric field through the region. Progressing from the drain toward the source, the channel becomes "less stressed" by the applied V_{ds} and eventually becomes un-pinched. From this point to the source the channel is in the low-field region [18]. Here, both drift and diffusion current components act. The saturation current model that follows includes the drift and diffusion components in the appropriate regions of the channel to obtain an expression for the current.

In the high-field region, the carriers are velocity saturated and the dominating current component is due to drift. The first step in establishing a saturation region drain current equation is to find the length of the low-field region, L_1. When the channel position, x, is equal to L_1, the electric field will reach its critical value and the carriers will become velocity saturated.

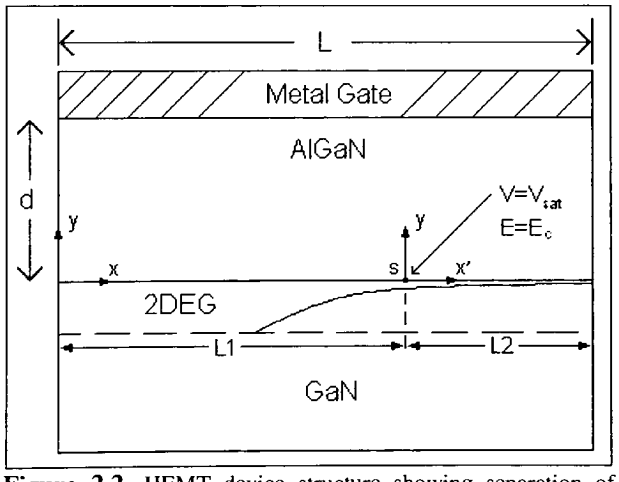

Figure 2.2. HEMT device structure showing separation of channel into low-field region, L_1, and high-field region, L_2. Point s is at the onset of saturation where the electric field assumes its critical value and the carrier velocity reaches its saturated value. Point s also acts as the origin for the offset axis y and x' [18].

The boundary condition just mentioned can be utilized once an expression is obtained for the channel potential and electric field in the low-field region. An expression for $V_c(x)$ is obtained by integrating equation (2.9) from 0 to x and is shown as,

$$\frac{b}{2}V_c^2(x)+\left(aI_{dsat}-bV_{gs}^{'}\right)V_c(x)+\left(bV_{gs}^{'}I_{dsat}R_s+xI_{dsat}-aI_{dsat}^2R_s-\frac{b}{2}I_{dsat}^2R_s^2\right)=0 \quad (2.14)$$

where

$$a=\frac{\mu_0 E_c-v_{sat}}{E_c v_{sat}}$$

$$b=\frac{W\mu_0\varepsilon(m)}{d}$$

The details of this derivation as well as the solution for $V_c(x)$ are shown in Appendix II. With an expression found for the channel potential, the electric field can be found by taking the derivative with respect to position, x, and, since the integration is starting at the source and progressing toward the drain which is at a higher potential, a negative sign is placed in front of the derivative as,

$$E(x)=-\frac{dV_c(x)}{dx}=\frac{I_{dsat}}{\sqrt{\left(aI_{dsat}-bV_{gs}^{'}\right)^2-2b(\alpha_0-\beta_0)}} \quad (2.15)$$

where

$$\alpha_0=xI_{dsat}+bV_{gs}^{'}I_{dsat}R_s$$

$$\beta_0=aI_{dsat}^2R_s+\frac{b}{2}I_{dsat}^2R_s^2$$

When $x = L_1$, $E(x) = -E_c$, and L_1 can be solved for using (2.15),

$$L_1=\frac{a^2I_{dsat}}{2b}+\frac{bV_{gs}^{'2}}{2I_{dsat}}+aI_{dsat}R_s+\frac{bI_{dsat}R_s^2}{2}-aV_{gs}^{'}-bV_{gs}^{'}R_s-\frac{I_{dsat}}{2bE_c^2} \quad (2.16)$$

With an expression obtained for the length of the low-field region, attention can now be focused on the high-field region ($L_1 \le x \le L$) where the Poisson equation must be used to account for current and field continuity. The 2-dimensional Poisson equation near the drain end of the channel is given as [18],

$$\frac{\partial^2 V_c(x',y)}{\partial x'^2}+\frac{\partial^2 V_c(x',y)}{\partial y^2}=-\frac{q}{\varepsilon(m)}N_D(y) \tag{2.17}$$

where x' is defined as follows [18]

$0 \le x' \le L_2 = L - L_1;$

$x' = x - L_1$

It is assumed that there is a uniform doping profile in the doped AlGaN layer. The following boundary conditions concerning the doping concentration result,

$N_D(y) = 0, \qquad 0 \le y \le d_i$

$N_D(y) = N_D, \qquad d_i \le y \le d = (d_d + d_i)$

Equation (2.17) is subject to the following boundary conditions [18]:

$$V_c(0,y) = (V_{gs} - \varphi_m(m)) + \frac{q}{\varepsilon(m)}\left(\int_0^d N_d(y)dy - \frac{I_{dsat}}{qWv_{sat}}\right)(d-y)$$
$$-\frac{q}{\varepsilon(m)}\int_y^d dy\int_y^d N_d(y)dy \tag{2.18a}$$

$$\frac{\partial V_c(0,0)}{\partial x'} = -E_c \tag{2.18b}$$

$$V_c(x',d) = (V_{gs} - \varphi_m(m)) \tag{2.18c}$$

$$\frac{\partial V_c(x',0)}{\partial y} = -\frac{qn_{sat}}{\varepsilon(m)} \tag{2.18d}$$

In equation (2.18a) the second term in the parentheses accounts for charge compensation in the channel. The net result of the expression inside the parentheses equals the number of electrons in the channel that did not come from the n-doped AlGaN layer. Following the solution outlined in [18], the solution to (2.17) is found by summing together the 1D Poisson equation solution and the 2D Laplace equation solution as,

$$V(x',y) = \Psi(y) + \Phi(x',y) \tag{2.19}$$

where

$$\frac{\partial^2 \Psi(y)}{\partial y^2} = -\frac{q}{\varepsilon(m)} N_d(y)$$

$$\frac{\partial^2 \Phi(x', y)}{\partial x'^2} + \frac{\partial^2 \Phi(x', y)}{\partial y^2} = 0$$

Using the boundary conditions listed in (2.18), a solution for the 1D Poisson equation in (2.19) is shown to be,

$$\Psi(y) = V_{gs} - \varphi_m(m) + \frac{q}{\varepsilon(m)}\left(\int_0^d N_d(y)dy - n_{sat}\right)(d-y) - \frac{q}{\varepsilon(m)}\int_y^d dy \int_y^d N_d(y)dy \quad (2.20)$$

Combining equations (2.18), (2.19), and (2.20) lead to the boundary conditions for the 2D Laplace equation as,

$$\Phi(0, y) = 0 \quad (2.21a)$$

$$\frac{\partial \Phi(0,0)}{\partial x'} = -E_c \quad (2.21b)$$

$$\Phi(x', d) = 0 \quad (2.21c)$$

$$\frac{\partial \Phi(x',0)}{\partial y} = 0 \quad (2.21d)$$

Using the method of separation of variables and the boundary conditions above, the solution for the 2D Laplace equation can readily be obtained as [18],

$$\Phi(x', y) = \frac{2dE_c}{\pi} \sinh\left(\frac{\pi}{2d}x'\right)\cos\left(\frac{\pi}{2d}y\right) \quad (2.22)$$

Adding together equations (2.20) and (2.22) yields the 2D channel potential in the high-field region,

$$\begin{aligned} V_c(x', y) = {} & \frac{2dE_c}{\pi} \sinh\left(\frac{\pi}{2d}x'\right)\cos\left(\frac{\pi}{2d}y\right) + (V_{gs} - \varphi_m(m)) \\ & + \frac{q}{\varepsilon(m)}\left(\int_0^d N_d(y)dy - \frac{I_{dsat}}{qWv_{sat}}\right)(d-y) \\ & - \frac{q}{\varepsilon(m)}\int_y^d dy \int_y^d N_d(y)dy \end{aligned} \quad (2.23)$$

Since y equals 0 at the hetero-interface, equation (2.23) simplifies to

$$V_c(x',y) = \frac{2dE_c}{\pi}\sinh\left(\frac{\pi}{2d}x'\right) + (V_{gs} - \varphi_m(m)) + \frac{q}{\varepsilon(m)}\left(\int_0^d N_d(y)dy - \frac{I_{dsat}}{qWv_{sat}}\right)d - \frac{q}{\varepsilon(m)}\int_0^d dy\int_0^d N_d(y)dy \quad (2.24)$$

Using equation (2.24), an expression can be obtained for the voltage dropped across the high-field region by evaluating equation (2.24) at either end of the high field region ($x' = L_2$, and $x' = 0$) and finding the difference. The resulting voltage across the high-field region is represented as,

$$V_H = \frac{2dE_c}{\pi}\sinh\left(\frac{\pi L_2}{2d}\right)$$

To obtain an expression for the saturation current, another equation must be formed that can represent the voltage across the high-field region. The following equation illustrates an alternate way for finding the high-field voltage drop:

$$V_H = V_c(x)\big|_{x=L} - V_c(x)\big|_{x=L_1}$$

The first term in the above V_H equation is the drain-to-source bias voltage and the second term can be found by plugging (2.16) into (2.14) and solving for $V_c(x)$. Equating both high-field region voltage equations and re-arranging some variables results in the following I_{dsat} expression,

$$0 = I_{dsat}^2\left(1 - a^2E_c^2 - 2abE_c^2R_s - b^2E_c^2R_s\right) + I_{dsat}\left(2bE_c^2L + 2abV_{gs}'E_c^2 + 2b^2V_{gs}'R_sE_c^2\right) - \left(bV_{gs}'E_c\right)^2 - \frac{4bdI_{dsat}E_c^2}{\pi}\sinh^{-1}\left(\frac{\pi}{2dE_c}\left(V_{ds} - V_{gs}' - I_{dsat}R_d + \frac{I_{dsat}}{b}\left(a + \frac{1}{E_c}\right)\right)\right) \quad (2.25)$$

The above equation can not be solved explicitly for I_{dsat} so an alternate approach is needed to obtain the saturation current for a given V_{gs} and V_{ds}. First, equation (2.25) must be re-arranged so that a solution can be realized. Bringing both negative terms to the other side of the equal sign gives

$$\left(bV_{gs}^{'}E_c\right)^2+\frac{4bdI_{dsat}E_c^2}{\pi}\sinh^{-1}\left(\frac{\pi}{2dE_c}\left(V_{ds}-V_{gs}^{'}-I_{dsat}R_d+\frac{I_{dsat}}{b}\left(a+\frac{1}{E_c}\right)\right)\right)=$$

$$I_{dsat}^2\left(1-a^2E_c^2-2abE_c^2R_s-b^2E_c^2R_s\right)+I_{dsat}\left(2bE_c^2L+2abV_{gs}^{'}E_c^2+2b^2V_{gs}^{'}R_sE_c^2\right)$$

Now that (2.25) has been formatted to an "$m = n$" format, both sides can be plotted individually as a function of V_{gs}, V_{ds}, and $I_{d,sat}$, and the intersection point between the two curves provides the values that can be used to generate an I_{ds} vs. V_{ds} plot.

In the following sections, equations (2.25) and (2.13) are used to derive device small-signal characteristics such as transconductance, drain conductance, and parasitic gate capacitances.

2.7 Transconductance, g_m, and Output Conductance, g_0

Transconductance, g_m, is an expression of the performance of a transistor. In general, the larger the transconductance of a device, the greater the gain (amplification) it is capable of delivering, when all other factors are held constant. The transconductance of any transistor is defined as the magnitude of change in drain current for a given change in gate-to-source voltage while keeping the drain-to-source voltage constant. g_m can be viewed as a sensitivity parameter relating the input voltage to the output current; the higher the g_m value, the greater the change in output current for a change in input voltage. The mathematical equation can be found by taking the derivative of the drain current with respect to the input voltage (V_{gs} in this case),

$$g_m=\left.\frac{\partial I_d}{\partial V_{gs}}\right|_{V_{ds}} \tag{2.26}$$

The transconductance was derived for both regions of operation, linear and saturation. The result for the linear region is,

$$g_{m,lin}=\left.\frac{\partial I_{d,lin}}{\partial V_{gs}}\right|_{V_{ds}}=\frac{I_{d,lin}b\left(R_d+R_s\right)-bV_{ds}}{2\alpha I_{d,lin}+\beta} \tag{2.27}$$

where α, β, a, and b have the same values as they did in the linear region drain current equation. The transconductance in the saturation region was found to be

$$g_{m,sat} = \left.\frac{\partial I_{dsat}}{\partial V_{gs}}\right|_{V_{ds}} = -\frac{\alpha_5 + 2bI_{dsat}E_c\alpha_4}{2I_{dsat}\alpha_1 + \alpha_2 - \frac{4bdE_c^2\alpha_3}{\pi} + 2bI_{dsat}E_c\alpha_4\left(R_d - \frac{1}{b}\left(a + \frac{1}{E_C}\right)\right)} \quad (2.28)$$

where

$$\alpha_1 = 1 - a^2E_c^2 - 2abE_c^2R_s - b^2E_c^2R_s$$

$$\alpha_2 = 2b\dot{E}_c^2L + 2abV_{gs}^{'}E_c^2 + 2b^2V_{gs}^{'}R_sE_c^2$$

$$\alpha_3 = \sinh^{-1}\left(\frac{\pi}{2dE_c}\alpha_6\right)$$

$$\alpha_4 = \left(1 + \left(\frac{\pi}{2dE_c}\right)^2\alpha_6^2\right)^{-1/2}$$

$$\alpha_5 = I_{dsat}\left(2abE_c^2 + 2b^2R_sE_c^2\right) - 2b^2E_c^2V_{gs}^{'}$$

$$\alpha_6 = V_{ds} - V_{gs}^{'} - I_{dsat}R_d + \frac{I_{dsat}}{b}\left(a + \frac{1}{E_c}\right)$$

The output conductance, g_o, of a transistor is similar to the transconductance except that the incremental change in output current is observed with respect to an incremental change in the drain-to-source voltage while keeping the gate-to-source voltage constant. Unlike g_m, however, it is desirable to have as low a value as possible for g_o. This is important for high-power amplifier circuits where it is desirable to have the output depend on the input exclusively. The output conductance can be represented in terms of the output current, output voltage, and input voltage as,

$$g_o = \left.\frac{\partial I_d}{\partial V_{ds}}\right|_{V_{gs}} \quad (2.29)$$

g_o was derived for both regions of operation, linear and saturation. The result for the linear region is

$$g_{o,lin} = \left.\frac{\partial I_{d,lin}}{\partial V_{ds}}\right|_{V_{gs}} = \frac{I_{d,lin}\left(a - bR_d\right) + b\left(V_{ds} - V_{gs}^{'}\right)}{2\alpha I_{d,lin} + \beta} \quad (2.30)$$

The drain conductance for the saturation region is

$$g_{o,sat} = \left.\frac{\partial I_{dsat}}{\partial V_{ds}}\right|_{V_{gs}} = \frac{2bE_c I_{dsat}\alpha_4}{2I_{dsat}\alpha_1 + \alpha_2 - \frac{4bdE_c^2\alpha_3}{\pi} + 2bE_c I_{dsat}R_d\alpha_4 - 2E_c I_{dsat}\alpha_4\left(a + \frac{1}{E_c}\right)} \tag{2.31}$$

2.8 C_{gs} and C_{gd}

Of all the parasitic capacitances inherent to a device, the two most detrimental to a transistor's high-frequency response are arguably those that lie between the gate and source, and the gate and drain terminals. Either one can be modeled by determining the amount of charge that changes in the channel when an incremental change in potential occurs between the gate and source or gate and drain terminals. Equations (2.32) and (2.33) show how both capacitances can be found mathematically,

$$C_{gs} = \left|\left.\frac{\partial Q}{\partial V_{gs}}\right|_{V_{ds}}\right| \tag{2.32}$$

$$C_{gd} = \left|\left.\frac{\partial Q}{\partial V_{ds}}\right|_{V_{gs}}\right| \tag{2.33}$$

where Q is the amount of channel charge and can be found by integrating the sheet carrier density over the length of the channel and multiplying the result by the width of the channel and by the electron charge, q [22]. Since the two regions of operation under scrutiny are linear and saturation, the channel charge must also be found for both regions before the capacitance in each region can be determined. To find the total amount of charge in the channel in the linear region, equation (2.34) must be used,

$$-Q_{lin} = qW\int_0^L n_s(m,x)dx \tag{2.34}$$

Equation (2.11) is used for $n_s(m,x)$ since it includes the channel potential term. After performing the necessary integration and multiplication, the total channel charge in the linear region was found to be,

$$-Q_{lin} = \frac{W\varepsilon(m)L}{d}V_{gt} - \frac{W\varepsilon(m)}{d}\left(AL + \frac{2}{3C}(B-CL)^{3/2} - \frac{2}{3C}B^{3/2}\right) \tag{2.35}$$

where

$$A = \frac{bV_{gs}^{'} - aI_{d,lin}}{b}$$

$$B = \frac{\left(aI_{d,lin} - bV_{gs}^{'}\right)^2 - 2b^2V_{gs}^{'}I_{d,lin}R_s + 2abR_sI_{d,lin}^2 + \left(bI_{d,lin}R_s\right)^2}{b^2}$$

$$C = \frac{2I_{d,lin}}{b}$$

$$V_{gt} = V_{gs} - V_{th}(m)$$

Taking the derivative with respect to the gate voltage yields the expression for the gate-to-source capacitance, C_{gs}, in the linear region,

$$C_{gs,lin} = \left|\frac{\partial Q_{lin}}{\partial V_{gs}}\right|_{V_{ds}} = \frac{W\varepsilon(m)L}{d} - \frac{W\varepsilon(m)}{d} \cdot \left[L\frac{\partial A}{\partial V_{gs}} + \frac{2}{3}\frac{1.5C(B-CL)^{1/2}\left(\frac{\partial B}{\partial V_{gs}} - L\frac{\partial C}{\partial V_{gs}}\right) - (B-CL)^{3/2}\frac{\partial C}{\partial V_{gs}}}{C^2} - \frac{2}{3}\frac{1.5CB^{1/2}\frac{\partial B}{\partial V_{gs}} - B^{3/2}\frac{\partial C}{\partial V_{gs}}}{C^2} \right] \quad (2.36)$$

To find the gate-to-source capacitance in the saturation region, the total channel charge must first be found. To do this, the channel must once again be split into low-field and high-field regions and the charge for each region must be summed together to determine the total amount in the channel [22],

$$-Q_{sat} = qW\int_0^{L_1} n_s(m,x)dx + qW\int_{L_1}^{L} n_s(m,x)dx \quad (2.37)$$

The first term accounts for the total amount of charge in the low-field region where the gradual-channel approximation still holds and the second term accounts for the total amount of charge in the saturated (pinched-off) region. If the current density equation in the saturation region is represented as,

$$I_{d,sat} = qWv_{sat}n_{sat}(m,x)$$

then an expression for the carrier density in the saturation region can be obtained. Substituting in for the carrier densities in both integrals,

$$-Q_{sat} = qW\int_0^{L_1}\frac{\varepsilon(m)}{q(d_d + d_i + \Delta d)}\left(V_{gs} - V_{th}(m) - V_c(x)\right)dx + qW\int_{L_1}^{L}\frac{I_{dsat}}{qWv_{sat}}dx$$

and performing the necessary integration and multiplication yields the total channel charge in the saturation region as

$$-Q_{sat} = \frac{I_{d.sat}(L - L_1)}{v_{sat}} + \frac{WL_1\varepsilon(m)V_{gt}}{d} - \frac{W\varepsilon(m)}{d}\left(AL_1 + \frac{2}{3C}(B - CL_1)^{3/2} - \frac{2}{3C}B^{3/2}\right) \quad (2.38)$$

Taking the derivative of (2.38) with respect to the gate voltage,

$$C_{gs.sat} = \left|\frac{\partial Q_{sat}}{\partial V_{gs}}\right|_{V_{ds}}\Bigg| = \frac{g_m L}{v_{sat}} - \left(\frac{g_m L_1}{v_{sat}} + \frac{I_{d.sat}}{v_{sat}}\frac{\partial L_1}{\partial V_{gs}}\right) + \left(\frac{WL_1\varepsilon(m)}{d} + \frac{W\varepsilon(m)V_{gt}}{d}\frac{\partial L_1}{\partial V_{gs}}\right)$$
$$- \frac{W\varepsilon(m)}{d}\left(A\frac{\partial L_1}{\partial V_{gs}} + L_1\frac{\partial A}{\partial V_{gs}} + A' - B'\right) \quad (2.39)$$

where

$$A' = \frac{2}{3C^2}\left[1.5C(B - CL_1)^{1/2}\left(\frac{\partial B}{\partial V_{gs}} - \left\{C\frac{\partial L_1}{\partial V_{gs}} + L_1\frac{\partial C}{\partial V_{gs}}\right\}\right) - (B - CL_1)^{3/2}\frac{\partial C}{\partial V_{gs}}\right]$$

$$B' = \frac{2}{3C^2}\left[1.5CB^{1/2}\frac{\partial B}{\partial V_{gs}} - B^{3/2}\frac{\partial C}{\partial V_{gs}}\right]$$

To obtain expressions for the gate-to-drain capacitance, C_{gd}, the same approach can be done using the channel charge equations just derived for C_{gs}, except taking the derivative with respect to the drain voltage instead of the gate voltage in accordance with (2.33). Therefore, taking the derivative of (2.35) with respect to V_{ds} gives

$$C_{gd.lin} = \left|\frac{\partial Q_{lin}}{\partial V_{ds}}\right|_{Vgs}\Bigg| = \frac{W\varepsilon(m)}{d}\cdot$$
$$\left[L\frac{\partial A}{\partial V_{ds}} + \frac{2}{3}\frac{1.5C(B - CL)^{1/2}\left(\frac{\partial B}{\partial V_{ds}} - L\frac{\partial C}{\partial V_{ds}}\right) - (B - CL)^{3/2}\frac{\partial C}{\partial V_{ds}}}{C^2} - \frac{2}{3}\frac{1.5CB^{1/2}\frac{\partial B}{\partial V_{ds}} - B^{3/2}\frac{\partial C}{\partial V_{ds}}}{C^2}\right] \quad (2.40)$$

where

$$\frac{\partial A}{\partial V_{ds}} = -\frac{a}{b} g_{o,lin}$$

$$\frac{\partial B}{\partial V_{ds}} = \frac{1}{b^2}\left[2\left(aI_{d,lin} - bV_{gs}^{'}\right)\left(ag_{o,lin}\right) - 2b^2 V_{gs}^{'} R_s g_{o,lin} + 4abR_s I_{d,lin} g_{o,lin} + 2b^2 R_s^2 I_{d,lin} g_{o,lin}\right]$$

$$\frac{\partial C}{\partial V_{ds}} = \frac{2}{b} g_{o,lin}$$

Taking the derivative of (2.38) with respect to V_{ds} gives the gate-to-drain capacitance in the saturation region,

$$C_{gd,sat} = \left|\frac{\partial Q_{sat}}{\partial V_{ds}}\bigg|_{V_{gs}}\right| = \frac{Lg_{o,sat}}{v_{sat}} - \left[\frac{I_{dsat}}{v_{sat}}\frac{\partial L_1}{\partial V_{ds}} + \frac{L_1}{v_{sat}} g_{o,sat}\right] + \frac{W\varepsilon(m)V_{gt}}{d}\frac{\partial L_1}{\partial V_{ds}}$$
$$- \frac{W\varepsilon(m)}{d}\left[A\frac{\partial L_1}{\partial V_{ds}} + L_1\frac{\partial A}{\partial V_{ds}} + A'' - B''\right] \qquad (2.41)$$

where

$$A'' = \frac{2}{3C^2}\left[1.5C(B - CL_1)^{1/2}\left(\frac{\partial B}{\partial V_{ds}} - \left(C\frac{\partial L_1}{\partial V_{ds}} + L_1\frac{\partial C}{\partial V_{ds}}\right)\right) - (B - CL_1)^{3/2}\frac{\partial C}{\partial V_{ds}}\right]$$

$$B'' = \frac{2}{3C^2}\left[1.5CB^{1/2}\frac{\partial B}{\partial V_{ds}} - B^{3/2}\frac{\partial C}{\partial V_{ds}}\right]$$

2.9 Cutoff Frequency, f_T

The cutoff frequency of a transistor is a figure of merit used to characterize the switching speed of the transistor. Though this pertains more to digital circuitry than RF power circuits, it is also the frequency at which the forward gain $|h_{21}|$ of the intrinsic device becomes unity [1] which is applicable to high-frequency analog circuits. To accurately calculate this, the y-parameters must be found for the device and converted to h-parameters. The h-parameters are then used to calculate the f_T of the transistor. A simplified, though less accurate, method can also be used to calculate the cutoff frequency. As shown in [1], dividing

the intrinsic transconductance of the device by the total of the parasitic gate capacitance (C_{gs} plus C_{gd}), f_T can be determined. Equation (2.42) gives the equation,

$$f_T = \frac{g_m}{2\pi\left(C_{gs} + C_{gd}\right)} \tag{2.42}$$

The "2π" is included in the denominator to yield a result in terms of hertz and not radians. Often C_{gd} can be excluded from the expression altogether since it is usually orders of magnitudes less than C_{gs}. In an attempt to keep this crude simplification as accurate as possible, C_{gd} is kept in. Equation (2.42) is used to plot f_T as functions of V_{ds} and L and the resulting relationships will be presented and discussed.

Materials and Apparatus

3.1 Hardware Data

Published experimental results on numerous HEMT samples were collected and analyzed for comparison. Results reported by Wu *et al.* [23] provided the experimental data that is used in this thesis for comparison with calculated values. [23] was used largely because the proposed model in Chapter 2 closely follows an I/V model proposed by Chang *et al.* [22] and Rashmi *et al.* [18] in which [23] was used as a benchmark to confirm the validity of the model. As will be shown in the subsequent chapter, for any results that couldn't be compared with experimental data, multiple references were obtained to verify that both the curve profiles and curve magnitudes were okay.

The physical structure of the sample fabricated in [23] was grown by Metal Organic Chemical Vapor Deposition (MOCVD) on a C-plane sapphire substrate. Fig. 3.1 shows the device with appropriately labeled regions.

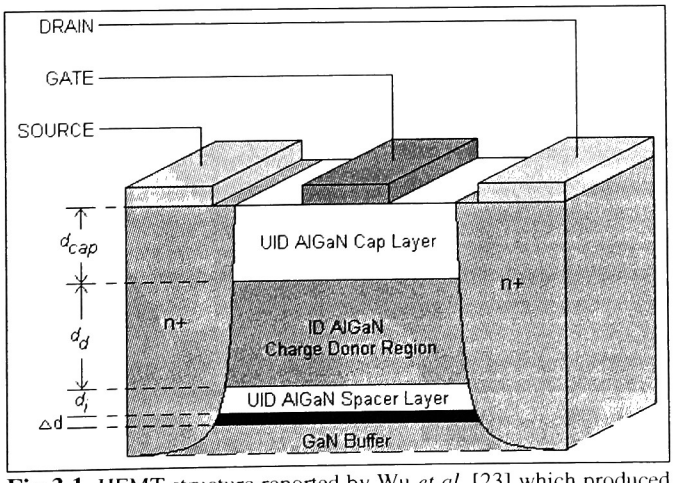

Fig 3.1. HEMT structure reported by Wu *et al.* [23] which produced the results used for comparison with the proposed model.

A 200 Å thick GaN nucleation layer was grown on top of the substrate followed by a 2 μm thick GaN insulation (buffer) layer. The $Al_{0.15}Ga_{0.85}N$ barrier grown between the GaN buffer and Schottky gate

consisted of a 30 Å unintentionally doped spacer layer, 220 Å n-doped charge donor layer, and a 150 Å cap layer. The 220 Å charge donor layer was silicon doped to $2x10^{18}$ cm^{-3} The source and drain ohmic contacts were annealed at 900 °C for 30 seconds and were made of Titanium/Aluminum/Nickel/Gold in proportions of 250 Å/2000 Å/400 Å/450 Å. Transfer contact resistance was measured to be in the range 0.5 to 0.7 Ω-mm. The device was grown to have a gate width of 75 μm, gate length of 1 μm, source-drain spacing of 3μm, and gate-drain spacing of 1 μm. The cap layer was grown in the structure to help suppress gate leakage currents by raising the HEMT peak barrier height [24,25]. "UID" and "ID" in the figure stand for unintentionally doped and intentionally doped, respectively. Table 3.1 lists all necessary device geometries and doping densities that were used in the subsequent chapter to obtain results. The 2DEG distance from the interface, Δd, was assumed to be 40 Å for all calculations.

Table 3.1. Device parameters as reported by Wu *et al.* [23] that were used for simulations.

Structure Parameter	**Value**
AlGaN Cap Layer Thickness, d_{cap}	150 Å
AlGaN Charge Donor Layer Thickness, d_d	220 Å
AlGaN Spacer Layer Thickness, d_i	30 Å
2DEG Electron Cloud Distance from Interface, Δd	40 Å
GaN Buffer Layer Thickness	2 μm
Doping Concentration of Charge Donor Region, N_D	$2x10^{18}$ cm^{-3}
Gate Width, W	75 μm
Gate Length, L	1 μm
Drain-to-Source Spacing, L_{DS}	3 μm
Gate-to-Drain Spacing, L_{GD}	1 μm

This chapter aimed to give the reader an understanding of the materials/resources necessary to complete this thesis work. Since no hardware tests were performed, all simulation results were compared against reported experimental data.

Chapter 4

Results and Discussion

4.1 Introduction

All results shown in this section were obtained using the equations derived in Chapter 2 – Theory / Model Formulation and are compared with experimental data when possible. In this work, experimental data has been collected from published results to verify the model. Numerous sources citing simulated data are used to substantiate any graphs that could not be reinforced with experimental results. The layout of this chapter follows the chapter 2 layout closely in that the charge control of the 2DEG will be analyzed first, followed by the current-voltage results, and finally small-signal parameter results.

All calculations were done using the material constants given in Table 4.1 and material equations given in Table 4.2.

Table 4.1. Table of constants used for results calculations with values and references.

Constant Description	Value	Reference
Electron Rest Mass, m_0	9.1094×10^{-31} kg	[26]
GaN Electron Mass, m^*	$0.22m_0$ kg	[20]
AlN Electron Mass, $m_{e,AlN}$	$0.33m_0$ kg	[20]
Permittivity constant, ε_0	8.85418×10^{-14} F/cm	[26]
GaN Dielectric constant, ε_{GaN}	$10\varepsilon_0$ F/cm	[27]
AlN Dielectric constant, ε_{AlN}	$8.5\varepsilon_0$ F/cm	[27]
Electron charge, q	1.602×10^{-19} C	[26]
Boltzmann constant, k_B	1.38066×10^{-23} J/K	[26]
Planck constant, h	6.62607×10^{-34} J·sec	[26]
GaN Lattice constant, a_{GaN}	3.189×10^{-10} m	[20]
AlN Lattice Constant, a_{AlN}	3.112×10^{-10} m	[20]
GaN Band Gap, $E_{g,GaN}$	3.42 eV	[20]
AlN Band Gap, $E_{g,AlN}$	6.13 eV	[20]

For constants that are given for GaN and AlN, Vegard's Law was used to extrapolate an equation for the AlGaN material where the mole fraction lies between 0 and 1. A few equations couldn't be formed using Vegard's Law due to the non-linear relationship between the material parameter and the mole fraction (such as the AlGaN band gap), so references were obtained that contained the necessary equations. Table 4.2 shows the equations that were formed using Vegard's Law and provides references for those that were looked up.

Table 4.2. AlGaN/GaN material system equations used to calculate results. m is the aluminum mole fraction. Room temperature (300K) was assumed for all calculations.

Equation Description	Equation	Reference
Schottky Barrier Height	$\varphi_m(m) = 0.91 + 2.44m$ eV	[16]
AlGaN Effective Electron Mass	$m_{e,AlGaN}(m) = (0.22 + 0.11m)m_0$ kg	Vegard's Law
AlGaN Dielectric constant	$\varepsilon(m) = (10 - 1.5m)\varepsilon_0$ F/cm	Vegard's Law
AlGaN Lattice constant	$a_{AlGaN}(m) = (3.189 - 0.077m) \times 10^{-10}$ m	Vegard's Law
AlGaN Energy Gap	$E_{g,AlGaN}(m) = 6.13m + 3.42(1-m) - m(1-m)$ eV	[20]
Conduction Band Discontinuity	$\Delta E_c = 0.7 * [E_{g,AlGaN} - E_{g,GaN}]$	[20]
Thermal Voltage	$V_T = k_B T/q$	-

4.2 Sheet Carrier Density, n_s, vs. Fermi Level, E_f

In Section 2.1, a relationship between channel density and Fermi level was derived and a procedure presented on how to obtain a plot for n_s vs. E_f. Equation (2.4) was used to produce the results shown in Fig. 4.1 where the sheet carrier density is plotted as a function of the Fermi level. A temperature of 300K was used and E_f was varied from $-6V_T$ to $3V_T$ to produce the results.

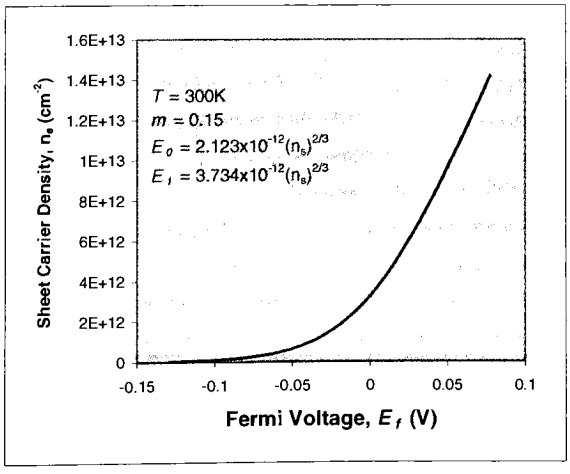

Fig 4.1. The channel sheet carrier density as a function of E_f at for T = 300K. E_f varied from $-6V_T$ to $3V_T$

It is clear from the figure that there exists a direct relationship between n_s and E_f, though the type of relationship depends on which section of the curve is being analyzed. At Fermi voltages less than ~-0.025 V, n_s is observed to have an exponential dependence on E_f which turns to a linear dependence when E_f is increased from there. The exponential dependence of n_s on E_f is difficult to see at Fermi voltages less than -0.1 V due to the magnitude of n_s being very small relative to the plotted magnitude. This type of dependence is more evident in Fig. 4.2 where n_s is plotted as a function of E_f on a logarithmic scale.

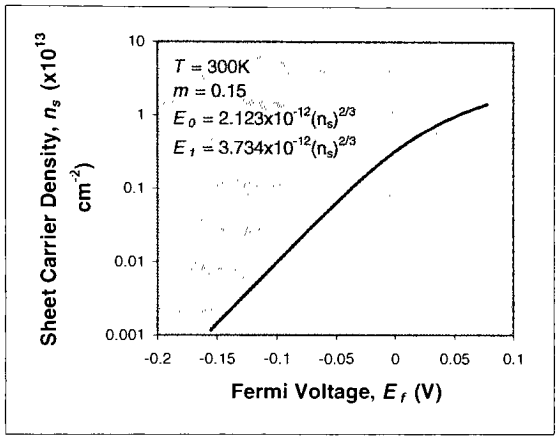

Fig 4.2. Sheet Carrier Density plotted logarithmically as a function of Fermi voltage to emphasize dependence type.

The exponential dependence of n_s on E_f is evident in Fig. 4.2 by the linear slope of the curve prior to $E_f \approx$ -0.025 V after which the curve begins to "slope over" corresponding to the linear portion of the Fig. 4.1 curve. This "slope over" indicates the point where charge control occurs between V_{gs} and n_s which will be covered in more detail in the subsequent section. The profile and magnitude of the Fig. 4.1 and Fig. 4.2 curves positively reinforce what was established in Section 2.1, which was the premise that as the Fermi level increases, the quantum well eigenstates become filled with electrons, raising the value of the sheet carrier density. The curves, furthermore, provide a visual interpretation to what equation (2.4) was providing mathematically. Although experimental values could not be obtained for comparison, several references have been gathered that report similar n_s-E_f behavior [1,12,28]. The cited references help legitimize the use of equation (2.4) for predicting n_s values.

Fig. 4.2, in addition to showing the n_s dependence more effectively than Fig. 4.1, is also better used for reading values. In the exponential region (E_f < -0.025 V), n_s increased from 1.435x10^{10} cm^{-2} to 1.47x10^{12} cm^{-2} as E_f increased in value from -0.15 V to -0.0258 V. As E_f enters the "linear region" voltage range (E_f > -0.025 V), n_s is observed to have less response to increases in E_f, reaching a maximum value of 1.42x10^{13} cm^{-2} for an E_f of 0.0776 V.

A relationship between the 2DEG concentration, n_s, and the Fermi level, E_f, has been established in this section. For E_f values less than $-V_T$, n_s exhibits an exponential dependence that turns to a linear dependence once E_f exceeds $-V_T$. The next section shows how both depend on V_{gs} and how gate capacitance can be determined from a plot of n_s vs. V_{gs}.

4.3 Sheet Carrier Density, n_s, vs. Gate Voltage, V_{gs}

In Section 2.1, equation (2.8) was derived relating the sheet carrier density to the gate voltage. Using that equation and equation (2.5), Fig. 4.3 was created to visually show the relationship. E_f is also shown on graph since it is also dependent on the gate voltage. Before analyzing the results shown in Fig. 4.3, some discussion is given to provide some understanding to the relationships between n_s, E_f, and V_{gs}. Looking at the energy band diagram provided in Chapter 1, it can be seen that the Fermi level aligns between all the materials in the system under equilibrium conditions. When a gate voltage is applied, the conduction band discontinuity shifts accordingly (down for positive V_{gs} values, up for negative V_{gs} values) and the difference in potential between the bottom of the conduction band discontinuity and the Fermi level changes. At gate voltages less than threshold, the Fermi level lies well below the bottom of the quantum well and the concentration of electrons occupying the discrete eigenstates is very low (10^6 cm^{-2} ~ 10^{10} cm^{-2}). As V_{gs} approaches threshold, E_f is expected to be close to the bottom of the well and the charge density in the channel is expected to become large. Above threshold, the gate has complete charge control over the channel density and a direct linear relationship results. This region is known as the strong inversion region and is the desirable operating region for amplifier applications.

Equation (2.6) was used to calculate the threshold voltage, V_{th}, for the given sample (m = 0.15, d_d = 220 Å, d_i = 30 Å, Δd = 30 Å, N_D = 2x10^{18} cm^{-3}) which was found to be -3.218 V. This was the calculated

threshold voltage at V_{gs} = 0 V and assuming a 90% ionization of the donor impurity atoms (an assumption treated as a fitting parameter). If V_{gs} does not equal 0 V, the threshold voltage will vary slightly about this value but by no more than a few tenths of a volt depending on the location of the Fermi level [1]. The results shown in Fig. 4.3 can now be observed with a bit more understanding. At low V_{gs} values (≤ -3.5 V), the Fermi level is well below the bottom of the conduction band edge and eigenstates. Naturally, the 2DEG concentration is very small and in fact the magnitude of the 2DEG concentration is so small it can not be determined from inspection of the plot. The calculated value is on the order of 10^7. As V_{gs} approaches threshold, E_f begins to show signs of an exponential dependence on V_{gs} and n_s becomes large enough to be noticed on the graph. Increasing V_{gs} further results in n_s and E_f both having a linear dependence on V_{gs} showing the charge-control ability of the Schottky gate. This behavior is expected since equation (2.5) is inherently a linear equation relating n_s to V_{gs}. At V_{gs} = -3 V, n_s was calculated to be 4.96x10^{11} cm^{-2} and E_f to be -0.0569 V. At V_{gs} = 0 V, n_s was calculated to be 6.21x10^{12} cm^{-2} and E_f to be 0.0259 V. Increasing V_{gs} further to the maximum value shown on the figure (V_{gs} = 0.72 V), results in an n_s value of 7.53x10^{12} cm^{-2}. If V_{gs} were to increase even more, n_s would eventually attain a value on the order of 10^{13} cm^{-2}. It is this high order of 2DEG concentration that makes the AlGaN/GaN heterostructure system very attractive and, considered by most, superior to its AlGaAs/GaAs counterpart where concentration values rarely get above 10^{11} cm^{-2}

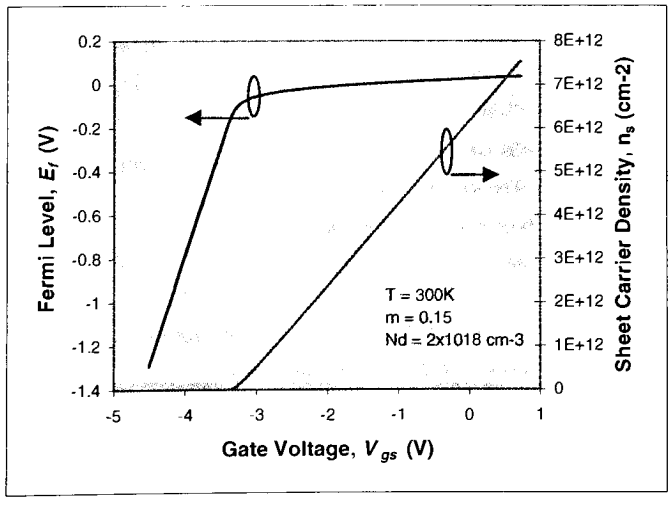

Fig 4.3. Fermi level and sheet carrier density plotted as a function of gate voltage. T = 300K, m = 15%, V_{ds} = 0 V.

The behavior of the curve in Fig. 4.3 has been demonstrated by others [15,19] as well. A closer look at these references reveals slightly different curve behavior at V_{gs} values close to threshold. In Fig. 4.3, the n_s curve is observed to very linearly with V_{gs} except where they meet at the x-axis. At this point, the exponential relationship is visible for very low values of n_s. In references [15] and [19] the exponential relationship between n_s and V_{gs} is more pronounced for V_{gs} values at threshold and below. The reason why the exponential relationship it is not visible with this model is due to the fact that only two eigenstates were considered in the quantum well. This limits the effectiveness of the model in predicting 2DEG concentration values to the strong inversion region only. To improve the precision of the model, more eigenstates need to be included, though the increase in eigenstates is important only if subthreshold concentration levels need to be known. More states did not need to be included here since the device is biased in the strong inversion region for all current-voltage calculations and small-signal calculations.

Since Fig. 4.3 is plotting charge vs. voltage, one would be able to extrapolate the capacitance of the device from the curve at a given gate voltage. Despite n_s not being visible at low V_{gs} values, its behavior was determined to be exponential with respect to V_{gs}. Therefore it could be said that the device's capacitance increases as V_{gs} increases and eventually saturates when V_{gs} becomes greater than V_{th}. Taking the slope of the n_s line, multiplying by the electron charge, q, the gate width, W, and gate length, L,

$$C_g = qWL\frac{\Delta n_s}{\Delta V_{gs}} = (1.602\times10^{-19}C)(75\mu m)(1\mu m)\frac{\Delta n_s}{\Delta V_{gs}}$$

revealed a gate capacitance of 228.55 fF. The gate capacitance is directly related to the distance between the gate and the conducting channel. Therefore, if the barrier thickness were increased, one could expect the slope of the curve to change by becoming smaller. One could also expect the threshold voltage to become increasingly negative to account for the extra electrons in the barrier, making it easier to turn the device on and more difficult to turn it off.

This section extended the relationship established in Section 2.1 to include V_{gs}. Both E_f and n_s are shown to be dependent on V_{gs}. For V_{gs} values at threshold and below, n_s has an exponential dependence that

turns linear for V_{gs} values at and above threshold. Plotting n_s vs. V_{gs} can lead to a calculation of the gate capacitance under zero drain bias conditions.

4.4 Sheet Carrier Density, n_s, vs. Barrier Thickness, d_d

Equation (2.5) was used to produce a plot of sheet carrier density, n_s, vs. AlGaN barrier layer thickness, d_d. Fig. 4.4 shows the aforementioned plot assuming $V_{gs} = E_f = V_{ds} = 0$ V. The graph shows what was briefly touched upon in the previous section concerning the relationship between the carrier density and the barrier thickness. Intuitively, one would expect the carrier density to increase as d_d increases for the simple reason that as d_d increases, a larger dopant area is "donating" electrons to the channel, increasing the channel density. From a quantitative standpoint, a barrier layer of thickness 250 Å doped at $1x10^{19}$ cm^{-3}, will not physically contain the same number of "depletable" electrons as a barrier, say, of thickness 350 Å doped at the same concentration. The 350 Å barrier has more room for dopant electrons and can therefore donate more electrons when biased to do so. This concept is reinforced by Fig. 4.4. The carrier density is observed to increase at a decreasing rate eventually stabilizing by increasing linearly as the barrier thickness increases.

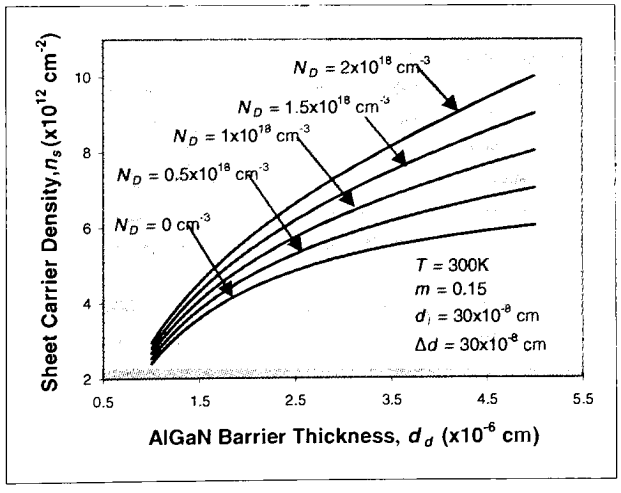

Fig 4.4. 2DEG density as a function of barrier thickness with doping density as a parameter. d_d varied between $100x10^{-8}$ cm and $500x10^{-8}$ cm. Doping concentration varied between 0 cm^{-3} and $2x10^{18}$ cm^{-3} in $0.5x10^{18}$ cm^{-3} increments. $V_{ds} = V_{gs} = 0V$.

As Section 4.3 would indicate, this type of 2DEG behavior is expected. A larger barrier thickness would increase the 2DEG density by increasing the amount of electrons available to the 2DEG. Fig. 4.4 also

indicates that the channel density is more sensitive to the barrier thickness if it is highly doped. For example, at any given barrier thickness between $100\text{x}10^{-8}$ cm and $500\text{x}10^{-8}$ cm, the higher doped barrier yields a higher 2DEG concentration than other barriers with lower doping concentrations. To quantify this concept, the curve corresponding to $N_D = 2\text{x}10^{18}$ cm^{-3} yields a sheet carrier density of $8.1428\text{x}10^{12}$ cm^{-2} at $d_d = 350\text{x}10^{-8}$ cm whereas the $N_D = 1\text{x}10^{18}$ cm^{-3} curve yields a carrier density of $6.8174\text{x}10^{12}$ cm^{-2} at $d_d = 350\text{x}10^{-8}$ cm. This makes sense since a higher doping concentration will yield a higher volume of available electrons.

An interesting result to note here is how the undoped curve behaves the same as those curves that are doped. This phenomenon is a current hot topic amongst researchers and has led many to fabricate and run tests on undoped barrier devices [29,30] with encouraging results. In fact, devices with undoped barriers have been observed to be less leaky, less noisy, and to yield higher electron mobilities [29].

Having established that both barrier thickness and doping density can affect the channel density, much can be said of their relationship to the high-frequency response of the device as well. Consider device "A" for example, which has barrier thickness $250\text{x}10^{-8}$ cm and a doping density of $1\text{x}10^{18}$ cm^{-3}. For comparison, consider device "B" which is in every way the same as device "A" except that its barrier thickness is $300\text{x}10^{-8}$ cm and its doping density is $0.5\text{x}10^{18}$ cm^{-3} According to Fig. 4.4, both yield doping densities of $5.73\text{x}10^{12}$ cm^{-2} This provides the designer with the certain degree of freedom when fabricating these devices. Should he/she decide to build device "A," he/she would enjoy the same amount of current drive as device "B," conserve vertical space on the wafer compared to device "B," but experience a degraded high frequency response compared to "B." This is because the gate capacitance will be higher due to the thinner barrier layer. Conversely, device "B" would provide an improved high frequency response but would take up more vertical space on the wafer. This type of tradeoff, space for frequency response, gives the designer the freedom to choose the optimum device geometry depending on the application. The foregoing capacitance tradeoff with the barrier thickness was discussed briefly by Rashmi *et al.* [15].

In this section, a relationship has been shown detailing carrier density dependence on barrier thickness. As d_d increases, n_s also increases as the number of free electrons in the barrier goes up.

4.5 Threshold Voltage, V_{th}, vs. Doping Density, N_D

Near the end of Section 4.3 it was stated how the number of electrons in the barrier directly affects the magnitude of the threshold voltage. A larger number of electrons affects V_{th} by making it more negative. It is therefore quite clear that either the doping density of the barrier layer or the barrier layer thickness would lead to an increasingly negative threshold voltage. Most fabricated HEMTs are inherently depletion-mode devices so increasing either d_d or N_D would make it even more difficult to turn them off. Most contemporary digital circuitry employs enhancement-mode FET devices which require a finite gate-to-source voltage to create a current channel connecting drain and source. This has led some GaN researchers to try the same with GaN-based HEMTs. Enhancement mode devices have been fabricated [31,32] and studied, although the fabrication process is far from being considered commercially viable, as is the case with any GaN-based circuit design.

Fig. 4.5 was constructed using equation (2.6) to show the V_{th} dependence on N_D with barrier thickness as a parameter. V_{th} is plotted as a function of N_D which is varied between $N_D = 0$ cm^{-3} to $N_D = 3.5\text{x}10^{18}$ cm^{-3} Five curves were made, each with a different barrier thickness (only thicknesses that are practical were considered in the simulation). The five thicknesses used were $d_d = 120\text{x}10^{-8}$ cm, $160\text{x}10^{-8}$ cm, $200\text{x}10^{-8}$ cm, $240\text{x}10^{-8}$ cm, and $280\text{x}10^{-8}$ cm.

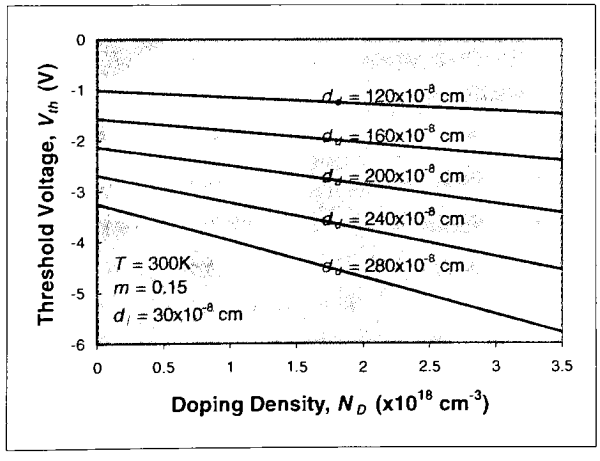

Fig 4.5. Threshold voltage, V_{th}, as a function of doping density, N_D, with barrier thickness, d_d, as a parameter. $V_{ds} = V_{gs} = 0$ V.

The graph shows that V_{th} becomes more negative as the doping density is increased. Furthermore, for a given doping density, V_{th} is more negative for thicker barrier layers. For example, a barrier of thickness

$160x10^{-8}$ cm at a doping density of $2x10^{18}$ cm^{-3} yields a V_{th} of -2.05 V whereas a barrier of $280x10^{-8}$ cm yields a V_{th} of -4.7 V. These results further reinforce the overlying concept which is that the number of electrons in the barrier help to dictate what the threshold voltage of the device will be. Barrier thickness and doping density are fabrication parameters used to control the number of free electrons in the barrier layer. As stated in Section 4.4, a designer has the freedom to choose which technique he/she will use to obtain a specified level of 2DEG density and threshold voltage. The results shown in Fig. 4.5 are similar to those reported by Rashmi *et al.* in [15].

4.6 Current-Voltage Characteristics

Before delving into current-voltage results, some background concerning the channel charge concentration, $|Q|$, versus drain voltage, V_{ds}, and channel carrier velocity, $v(x)$, versus drain voltage will be given to provide some insight into the I/V results. Equations (2.35) and (2.38) were used to observe the relationship between $|Q|$ and V_{ds}. Fig. 4.6 shows the channel charge density, Q, as a function of the applied drain-to-source bias with gate voltage as a parameter. Gate voltages of -2 V, -1 V, 0 V, and 1 V were used and V_{ds} was varied from 0 V to 6 V to obtain the results shown.

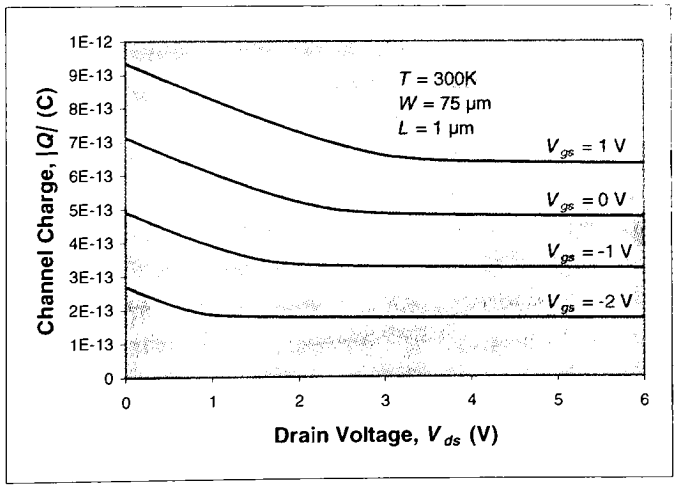

Fig 4.6. Channel charge as a function of V_{ds}. Gate voltages of -2 V, -1 V, 0 V, and 1 V were used for the curves. Width and length of device and temperature used in simulation are shown on the graph.

Each curve is observed to decrease linearly and eventually flatten out progressing to the right. The decrease in charge is expected by inspection of equation (2.11) which has the sheet carrier density, n_s, being

negatively related to the channel voltage. Since Q is directly related to n_s through the dimensions of the device and $V_c(x)$ is directly related to the magnitude of V_{ds}, it is clear that $|Q|$ should decrease as V_{ds} increases. The question now is, if $|Q|$ decreases as V_{ds} increases, then why does each curve experience a "flattening?" This is due to the device reaching the saturation region of operation. Like other short-channel FET devices, the current in the saturation region becomes "independent" of V_{ds} because the channel becomes pinched-off and the carriers are velocity saturated. In short-channel FETs, the drain current isn't completely independent of drain voltage in the saturation region because of channel-length modulation. This is evident by the slope that is still apparent in the "flat" part of the curves. With both the velocity of the carriers and the amount of current in the channel being constant, the amount of charge in the channel must also be constant. In the linear region this is not the case since the channel is not pinched-off nor are the carriers velocity saturated. In the linear region, the channel charge decreases linearly with a linear increase in $V_c(x)$. Despite the decrease in $|Q|$ observed for increasing drain biases, I_{ds} continues to increase because the carrier velocity increases rapidly with the lateral electric field. The velocity of the carriers in the channel increases rapidly as V_{ds} increases. With regards to the drain current, this increase in carrier velocity more than compensates for the decrease in channel charge and the drain current experiences an increase in magnitude. The results of the I-V curves can now be interpreted with more understanding.

Equations (2.13) and (2.25) were used to plot the I/V characteristic curves shown in Fig. 4.7. At first glance, the results resemble the I/V characteristics for most transistors where an initial linear increase in current eventually flattens out with increasing drain bias. As was done in Fig. 4.6, curves were obtained for gate voltages of -2 V, -1 V, 0 V, and 1 V. Good agreement was found between the predicted and measured values (Wu *et al.* [23]). All curves were predicted to within 8% of their measured values at all times. Drain currents of 537 mA/mm, 350 mA/mm, 189 mA/mm, and 65 mA/mm were calculated for gate voltages of 1 V, 0 V, -1 V, and -2 V, respectively, at a V_{ds} of 6 V. These match closely to the experimental values of 500 mA/mm, 350 mA/mm, 200 mA/mm, and 68 mA/mm at V_{ds} = 6 V.

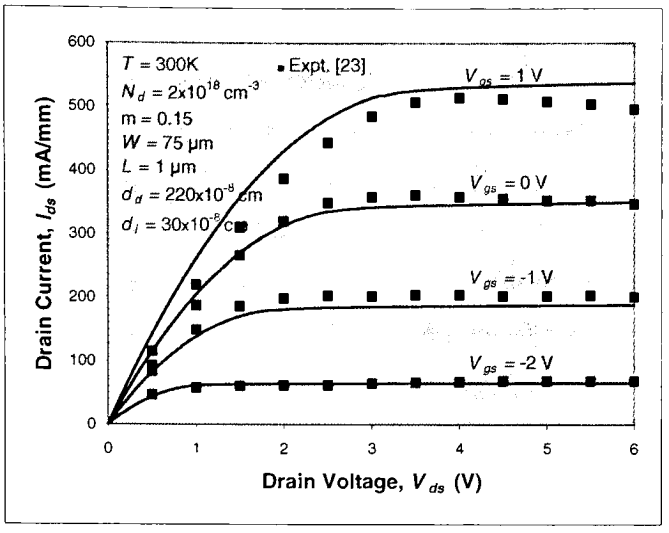

Fig 4.7. HEMT drain current as a function of drain voltage with gate voltage as a parameter. Current curves plotted for V_{gs} values of -2, -1, 0, and 1 V. Maximum current predicted was for V_{gs} = 1 V with an I_{ds} of 537 mA/mm. Solid lines correspond to model predictions; symbols correspond to experimental data [23].

A closer inspection of the experimental data corresponding to gate voltages of 1 V and 0 V reveals that the peak point in the data does not occur at the highest drain voltage of 6 V. Instead, experimental current was found to peak at drain voltages around 4 V and decrease slightly thereafter. One possible reason for this observed behavior is due to thermal effects. As the drain voltage increases, more current flows in the channel and the amount of power the transistor dissipates goes up. This dissipation of power causes the channel temperature to increase, subsequently decreasing the carrier mobility. The decrease in mobility results in a drain current decrease. Thermal effects were not, however, included in the current model and would be a valuable topic to research in the future. Comparing Figures 4.6 and 4.7, the relationship between the channel charge and the drain current is clear. Prior to saturation, all curves experience a linear change in magnitude. The change is positive for the current curves and negative for the charge curves. The length of the linear region is determined by the gate and drain voltages. Increasing the gate voltage means the transistor will remain in the linear region of operation for a larger V_{ds}. The higher gate voltage also results in there being an increase in channel charge, leading to the higher current levels.

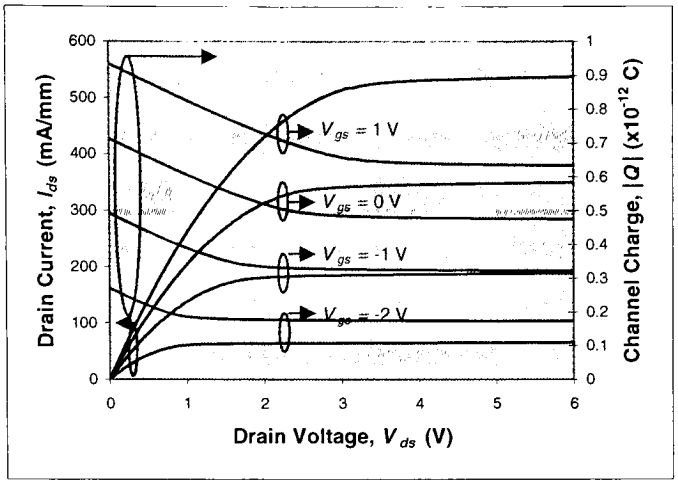

Fig 4.8. Drain current and channel charge both plotted as a function of drain voltage. Plot used to show relationship between each charge curve and its corresponding current curve in both linear and saturation regions of operation.

Fig. 4.8 plots both I_{ds} and Q against V_{gs} so the relationship can be appreciated visually. To increase the readability of the figure, the experimental data were not included.

Fig. 4.9 shows the relationship between the gate voltage and the drain current while keeping V_{ds} constant.

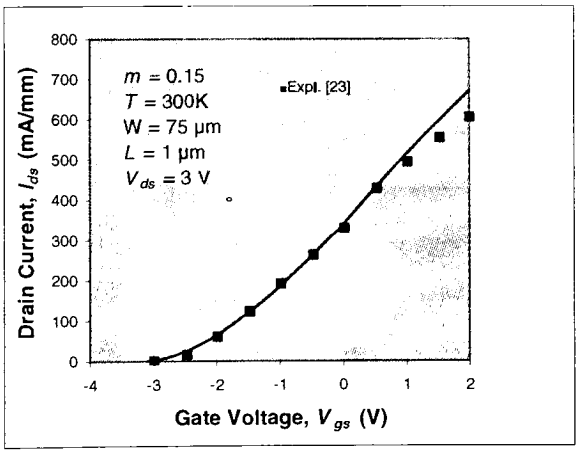

Fig 4.9. Drain current as a function of gate voltage. Temperature was set equal to 300K and V_{ds} to 3 V. Maximum I_{ds} of 671 mA/mm predicted against 604 mA/mm measured.

As V_{gs} increases from left to the right, the transistor is initially in saturation, eventually reaching linear operation around V_{gs} = 0 V (the transition can be seen on the graph as there is a slight discontinuity on the

curve around V_{gs} = -0.2 V). Excellent agreement was found between the calculated and measured values. The largest discrepancy occurred at V_{gs} = 2 V where the calculated value differed from experimental by 65 mA/mm. This difference is large but is still within 10% of the measured value. Making the discrepancy even less of an issue is the fact that the transistor will be used mostly for amplifier applications. This means that it will be biased in the saturation region so as to imitate a high-impedance current source. Agreement between the calculated and measured values in the saturation region (V_{gs} < 0 V) was found to be exceptional.

4.7 Transconductance, g_m, and Output Conductance, g_o

Equations (2.27) and (2.28) were used to calculate the transconductance, g_m. The results are shown in Fig. 4.10. The drain voltage was kept fixed at 2 V while the gate voltage was swept over a 6 V range starting at -4 V and ending at 2 V. Progressing from left to right, the curve is observed to increase sharply before leveling off and declining linearly. The sharp increase is due to the transistor entering the saturation region of operation immediately after V_{gs} becomes greater than V_{th}. As V_{gs} continues to increase, the transistor eventually transitions from the saturation region to the linear region resulting in a peak in the curve (this is seen around $V_{gs} \approx$ -0.7 V).

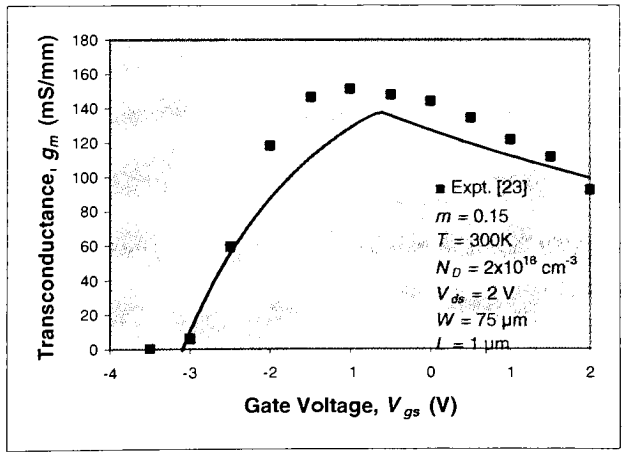

Fig 4.10. HEMT transconductance as a function of gate voltage. Drain voltage was kept fixed at 2 V and temperature at 300K. V_{gs} was varied from -4 V to 2 V.

As V_{gs} increases further, the transconductance is observed to decrease as the transistor extends into the linear region of operation. Reasonable agreement was found between predicted and measured values.

although the results here are not as encouraging as those obtained in Section 4.6. The largest discrepancies occur in the middle of the graph where the linear and saturation regions come together. The peak-calculated g_m value was found to be 137 mS/mm whereas the measured values went as high as 151 mS/mm. Since the HEMT will be biased to operate in the saturation region, the accuracy of the left half of the curve is of greater interest. The model accurately predicts g_m for low V_{gt} ($V_{gs} - V_{th}$) values. As V_{gt} becomes larger, the predicted values begin to deviate from the measured values, revealing that the proposed model is deficient for predicting transconductance values when the transistor is biased in saturation near the linear region. More research can be done to correct this problem and obtain better agreement between predicted and measured values throughout the specified V_{gs} range. Improvements in the mobility model or the charge control model could help to alleviate this problem.

Ideally, when a transistor is biased in the saturation region, the magnitude of the drain current should be independent of the drain voltage. This is because in saturation, the transistor behaves as a current source with infinite output resistance. Since the notion of infinite output resistance is impractical, the drain current is expected to increase to some degree with an increase in V_{ds} while the transistor is in saturation. To determine how the output current will change with a change in output voltage, it is imperative to know what the output conductance of the transistor is. Equations (2.30) and (2.31) were used to produce the output conductance, g_0, results shown in Fig. 4.11, and 4.12.

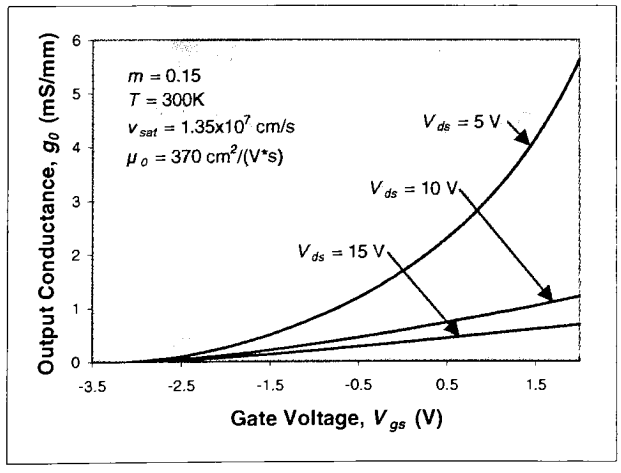

Fig 4.11. g_0 as a function of V_{gs} with V_{ds} as a parameter. V_{gs} varied from -3 V to 2 V. Curves were plotted for V_{ds} values of 5 V, 10 V, and 15 V.

Progressing from left to right, each curve in Fig. 4.11 is observed to increase linearly, eventually turning into an exponential increase (most obvious in the V_{ds} = 5 V curve). This type of behavior is expected as the transistor will initially be in the saturation region (meaning the output resistance will be high) making the transistor drain current resistant to changes in the drain voltage. This is evident from the figure as all curves are close in magnitude for V_{gs} values less than 0 V. As V_{gs} increases, the curves begin to separate from one another as they approach the linear region (note: for the V_{gs} range shown in the plot, none of the curves reach the linear region; higher drain voltages were used to represent high power scenarios, resulting in none of the curves reaching the linear regime for the V_{gs} voltages considered; V_{gs} was not increased further since gate voltages will not practically be that high.). The "V_{ds} = 5 V" curve is observed to increase the fastest, followed by the 10 V and 15 V curves. The curves increase in this order because as V_{ds} decreases, a lower V_{gs} value is needed to get the device into the linear region. In other words, a device under a drain-to-source bias of 5 V will reach the linear region of operation faster with increasing V_{gs} than the same device under a drain-to-source bias of 10 V. Since the lower drain bias device will reach the linear region at lower V_{gs} than the others, its output conductance will increase at a faster rate with respect to V_{gs}. At V_{gs} = 0 V, g_0 values of 1.67 mS/mm, 0.58 mS/mm, and 0.35 mS/mm were calculated for V_{ds} values of 5 V, 10 V, and 15 V, respectively. Increasing V_{gs} to 2 V resulted in an increase in each g_0 value to 5.63 mS/mm (V_{ds} = 5V), 1.21 mS/mm (V_{ds} = 10V), and 0.68 mS/mm (V_{ds} = 15 V), respectively.

It is also useful to analyze how the output conductance relates to the drain voltage. Plotting g_0 as a function of V_{ds} makes the concepts discussed in the previous paragraph easier to understand. Fig. 4.12 shows the relationship between g_0 and V_{ds}.

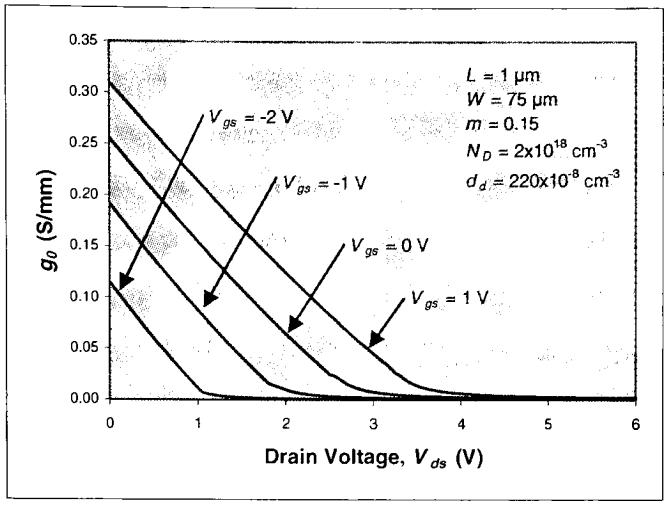

Fig 4.12. Output conductance, g_0, as a function of V_{ds} with V_{gs} as a parameter. Curves plotted for V_{gs} values of -2 V, -1 V, 0 V, and 1 V.

Fig. 4.12 closely resembles Fig. 4.6 in that both show curves following similar trends; a linear decrease eventually leveling off and becoming flat. The linear nature of the curves prior to reaching saturation is due to the transistor being biased in the linear region of operation. As V_{ds} increases, the transistor approaches the saturation region of operation and the output conductance concomitantly decreases further, reinforcing the fact that the transistor will behave like a current source in saturation and not like a resistor. In saturation, the curves are observed to behave as a current source showing little deviation with respect to V_{ds}. For a given V_{ds}, curves with higher V_{gs} values are observed to have a larger output conductance. This is because a higher V_{gs} value will result in a higher 2DEG concentration and a subsequent higher channel conductance. At no drain bias, V_{gs} values of -2 V, -1 V, 0 V, and 1 V yielded g_0 values of 0.11 S/mm, 0.19 S/mm, 0.26 S/mm, and 0.31 S/mm. The V_{gs} = -2 V curve was the first to reach saturation at a V_{ds} value of 1.05 V and with a transitioning g_0 of 6.4 mS/mm. The "transitioning g_0" is the value of the output conductance at the point where the saturation and linear regions meet. The other "transitioning g_0's" had values of 14.2 mS/mm, 24.9 mS/mm, and 28.4 mS/mm at V_{ds} values of 1.85 V, 2.5 V, and 3.25 V, respectively. At full drain bias, V_{ds} = 6 V, all g_0 values were calculated to be 0.002 mS/mm or less.

The output resistance, R_0, was plotted as a function of V_{ds} by taking the inverse of g_0 in Fig. 4.12. The results are shown in Fig. 4.13. R_0 is observed to increase to very large magnitudes of resistance as the biasing conditions extend into the saturation region of operation.

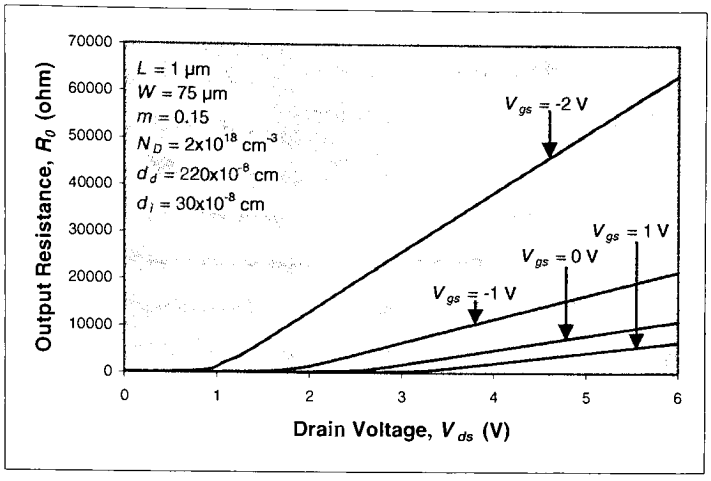

Fig 4.13. Output Resistance, R_0, as a function of Drain voltage with V_{gs} as a parameter. V_{gs} and V_{ds} values used for simulation are the same as those in Fig. 4.12.

For a given V_{ds} value, resistances corresponding to lower V_{gs} values are higher because the concentration of the 2DEG is lower consequently decreasing the conductance of the channel (shown in Fig. 4.12). At full drain bias (V_{ds} = 6 V) the calculated value of R_0 was found to be 6566 Ω, 11060 Ω, 21685 Ω, and 63621 Ω at V_{gs} values of 1 V, 0 V, -1 V, and -2 V, respectively. The magnitude of R_0 was also calculated for the linear region of each curve though the magnitude can not be appreciated by inspection of Fig. 4.13. Analysis of the plotted point values for the linear region was done to get some insight into the linear region resistance value since they can not be deduced from Fig. 4.13. Looking at the calculated data revealed that R_0 did not remain constant, though the variation pales in comparison to the saturation region values. For V_{gs} = -2 V, R_0 varied from 116 Ω to 1200 Ω though most of the points were below 700 Ω. For V_{gs} = -1 V, R_0 varied from 70 Ω to 830 Ω with most points coming below 400 Ω; and for V_{gs} = 0 V and 1 V, R_0 varied between 50 Ω ~ 550 Ω and 40 Ω ~ 350 Ω, respectively. With increasing gate voltage, the variation in linear region resistance decreased and became more constant, which is closer to the ideal (ideal scenario would be constant R_0 in the linear region). Multiple references have been found to help reinforce the results shown in this chapter. Besides [23] which reported the experimental data that was compared to the present model, [17,18,33,34] all report curve profiles similar to what has been shown here, in comparable magnitude ranges.

This section helped to establish relationships between the conductance parameters of the device and the terminal biasing voltages. When operating in the linear region, the device was shown to have high output conductance and low transconductance. This is consistent with a device behaving as a resistor and is similar to other FET devices in nature. Conversely, when in saturation, the device showed very low output conductance and a steep transconductance curve indicative of a device behaving as a current source.

4.8 C_{gs} and C_{gd}

Using equations (2.36) and (2.39) derived in Section 2.7, a plot of C_{gs} was obtained. Fig. 4.14 shows the results that were calculated using three different V_{ds} values (V_{ds} = 15 V, 10 V, and 5 V) over a V_{gs} range of -3.1 V to 2 V. The biasing conditions were chosen such that the capacitance of the device could be analyzed under circumstances that would be representative of real-world amplifier circuits. V_{gs} was not varied lower than -3.1 V since that would place the transistor in the subthreshold region and would provide un-reliable results. The proposed model does not provide any in-depth analysis of the subthreshold region which is why it has been avoided here. 2 V was used as the upper limit since it is representative of contemporary power supply voltage values. V_{ds} values as high as 15 V were used to show high-power capacitance values for the device. After an initial drop off, each curve is observed to increase at a decreasing rate, similar to a $y = x^{1/2}$ curve. The "V_{ds} = 5 V" curve is shown to have the highest C_{gs} value of the three curves, revealing that a lower drain voltage will result in higher C_{gs} values. It can be implied, then, that the transistor will operate better at higher frequencies at higher drain voltages. This characteristic favors the transistor as a high frequency power device.

This type of C/V response is expected by analyzing equation (2.11). This equation shows the 2DEG density, $n_s(m,x)$, being directly related to V_{gs} and negatively related to $V_c(x)$. Since $V_c(x)$ is directly related to V_{ds}, it can be said that $n_s(m,x)$ is negatively related to V_{ds} as well. This fact helps to explain the profile of the curves in Fig. 4.14 and why, for a given V_{gs} value, the curves with lower V_{ds} values yield higher C_{gs} values. It is clear from equation (2.11) that an increase in V_{ds} will result in a lower value for $n_s(m,x)$. A lower $n_s(m,x)$ value results in a lower value for the channel charge, Q.

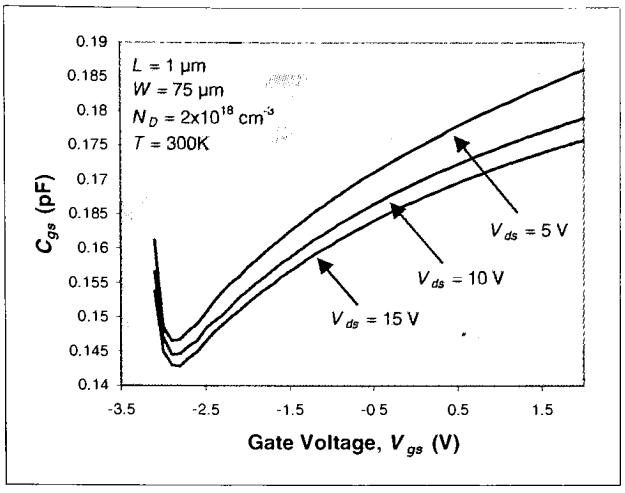

Fig 4.14. Gate-to-source capacitance as a function of gate voltage with drain bias as a parameter. Geometry of device is given on plot. High drain voltages used to demonstrate "high power" capacitance values.

Since $n_s(m,x)$ is more sensitive to changes in the gate voltage at lower V_{ds} values, the capacitance will be higher for lower V_{ds} values for a given V_{gs} value. This explains why the "V_{ds} = 15 V" curve is on the bottom and the "V_{ds} = 5 V" curve is on the top (though the difference in the magnitudes of the curves is small). In amplification applications, V_{gs} will assume a value between 0 V and the 2 V assumed here. Extrapolating data from the figure reveals C_{gs} values of 0.167 pF, 0.1696 pF, and 0.1747 pF for V_{ds} values of 15 V, 10 V, and 5 V, respectively, at V_{gs} = 0 V. Increasing V_{gs} to 2 V increased C_{gs} for each curve to values of 0.1758 pF (V_{ds} = 15 V), 0.179 pF (V_{ds} = 10 V), and 0.1861 pF (V_{ds} = 5 V).

This type of response has also been reported by [35,36]. In both, C_{gs} is shown to increase with V_{gs} but decrease with V_{ds} which is precisely the type of response that has been shown in Fig. 4.14. These results are encouraging and help to legitimize the model for small-signal parameter prediction.

Equations (2.40) and (2.41) were used to obtain the plot for C_{gd} vs. V_{gd} shown in Fig. 4.15. Curves were plotted for V_{gs} values of -1 V, 0 V, and 1 V. The biasing conditions considered here involved varying V_{gd} between -16 V and 1 V and V_{gs} between the values just mentioned. The conditions were once again chosen so that the "amplification capacitance" of the transistor could be analyzed. The figure shows each curve to have very small C_{gd} values (fF) at large negative V_{gd} values. Once V_{gd} exceeds -3 V, each curve is observed to increase dramatically and reach values in the pF range. This type of behavior once again supports the

argument that a GaN HEMT is an exceptional candidate for high-power, high-frequency applications. It can be seen from the figure that as V_{ds} rises, essentially making V_{gd} more negative, C_{gd} becomes smaller which is a desirable bi-product for the intrinsic frequency response of the device. Even though C_{gd} is the smaller of the two parasitic gate-to-channel capacitances (C_{gs} being the other), decreasing its value as much as possible only helps to further the high frequency performance of the transistor. Since only high drain voltages are of concern here, V_{gd} values larger than -3 V are ignored in the analysis. Pulling numbers from the plot reveals C_{gd} values of 6.7 fF, 9.9fF, and 17.2 fF for V_{gs} values of 1 V, 0 V, and -1 V, respectively, at a V_{gd} of -3 V. For higher drain voltages, both V_{gd} and C_{gd} become smaller. At V_{gd} = -10 V, C_{gd} had values of 0.725 fF, 0.66 fF, and 0.55 fF at V_{gs} values 1 V, 0 V, and -1 V, respectively. References [36,37] show curves similar to what has been shown here for C_{gd}.

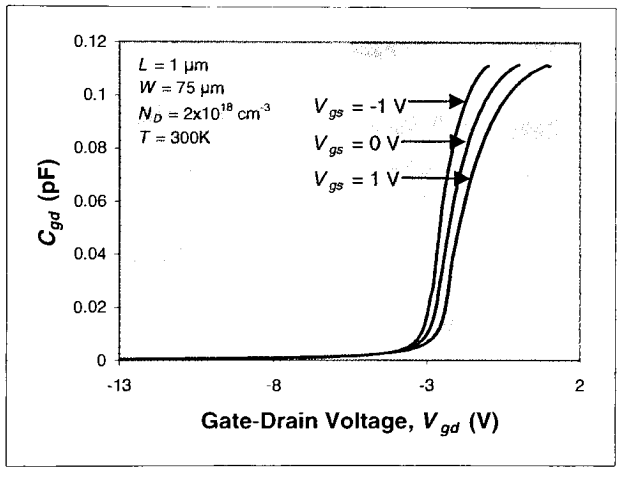

Fig 4.15. C_{gd} as a function of V_{gs} with V_{ds} as a parameter. All simulations were done for a 75μm x 1μm device with N_D = 2x10^{18} cm^{-3} and T = 300K. V_{ds} values of 15 V, 10 V, and 5 V were used.

The results for both C_{gd} and C_{gs} help to provide some insight into the frequency response of the device and how it is affected by the drain bias and gate length. Clearly from the discussion and results just provided an increase in V_{ds} helps to decrease both C_{gs} and C_{gd} which act as limiting agents for the frequency performance of the device.

4.9 Cutoff Frequency, f_T

argument that a GaN HEMT is an exceptional candidate for high-power, high-frequency applications. It can be seen from the figure that as V_{ds} rises, essentially making V_{gd} more negative, C_{gd} becomes smaller which is a desirable bi-product for the intrinsic frequency response of the device. Even though C_{gd} is the smaller of the two parasitic gate-to-channel capacitances (C_{gs} being the other), decreasing its value as much as possible only helps to further the high frequency performance of the transistor. Since only high drain voltages are of concern here, V_{gd} values larger than -3 V are ignored in the analysis. Pulling numbers from the plot reveals C_{gd} values of 6.7 fF, 9.9fF, and 17.2 fF for V_{gs} values of 1 V, 0 V, and -1 V, respectively, at a V_{gd} of -3 V. For higher drain voltages, both V_{gd} and C_{gd} become smaller. At V_{gd} = -10 V, C_{gd} had values of 0.725 fF, 0.66 fF, and 0.55 fF at V_{gs} values 1 V, 0 V, and -1 V, respectively. References [36,37] show curves similar to what has been shown here for C_{gd}.

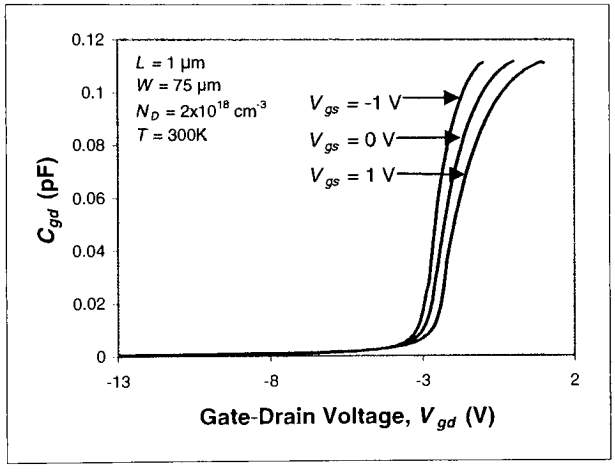

Fig 4.15. C_{gd} as a function of V_{gs} with V_{ds} as a parameter. All simulations were done for a 75µm x 1µm device with N_D = 2x10^{18} cm^{-3} and T = 300K. V_{ds} values of 15 V, 10 V, and 5 V were used.

The results for both C_{gd} and C_{gs} help to provide some insight into the frequency response of the device and how it is affected by the drain bias and gate length. Clearly from the discussion and results just provided an increase in V_{ds} helps to decrease both C_{gs} and C_{gd} which act as limiting agents for the frequency performance of the device.

4.9 Cutoff Frequency, f_T

Using equation (2.42) and varying the gate length, L, and drain voltage, V_{ds}, two figures were obtained for the cutoff frequency. Fig. 4.16 shows f_T as a function of V_{ds} and Fig. 4.17 shows f_T as a function of L. V_{ds} was varied from 0 V to 15 V and V_{gs} was set equal to 1 V for the plot of f_T vs. V_{ds}. Fig. 4.16 helps to reinforce what was established in Section 4.8 – the device frequency response improves with an increase in V_{ds}. This is due to the decrease in both C_{gd} and C_{gs} with the increase in V_{ds}. In can be seen from the figure that f_T flattens out for V_{ds} values in excess of 4 V indicating that both C_{gd} and C_{gs} also flatten out. Prior to "saturating" the curve is observed to increase quickly until V_{ds} reaches a value of about 4 V. This rapid increase in the value of f_T occurs because C_{gd} and C_{gs} are decreasing with the increase in V_{ds}. These results are encouraging for high power applications where V_{dg} ($V_{ds} - V_{gs}$), will typically have a magnitude of 3 V or higher. The region of interest lies to the right of V_{ds} = 4 V since this is the drain voltage that would make V_{dg} equal to 3 V (with V_{gs} = 1 V).

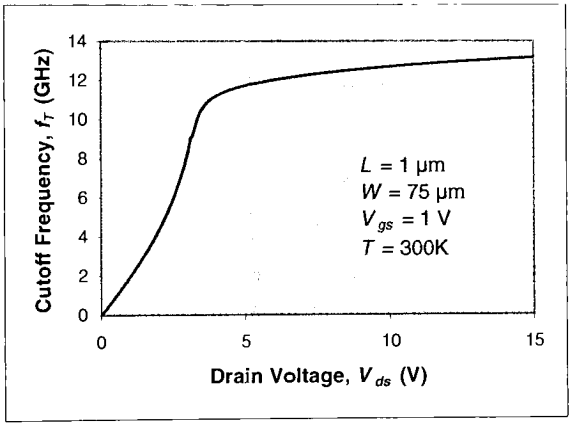

Fig 4.16. f_T as a function of V_{ds}. f_t calculated for device of width 75 μm and length 1 μm. V_{gs} of 1 V was used to represent typical gate amplification voltages.

Pulling numbers from the curve yielded f_T values of 11.19 GHz, 12.69 GHz, and 13.15 GHz for V_{ds} values of 4 V, 10 V, and 15 V, respectively. Similar curve profiles have also been reported by [38,39] where the device cut-off frequency has also been plotted as a function of drain voltage.

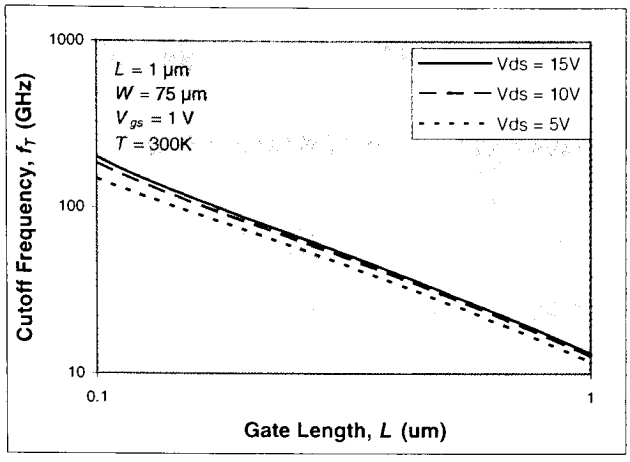

Fig 4.17. f_T as a function of gate length, L. Simulations done for drain biases of 15 V, 10 V, and 5 V. Dimensions of device were 75 μm by 1 μm.

Fig. 4.17 shows the relationship between the cutoff frequency, f_T, and the gate length, L. Curves were plotted for drain biases of 15 V, 10 V, and 5 V. All curves were plotted with a constant gate voltage of 1 V. The cutoff frequency is shown to have an inverse relationship with the gate length as f_T steadily decreases with the increase in L (note that both axes were converted to logarithmic type). Intuitively, this is expected since a larger L value would result in larger C_{gd} and C_{gs} values and a decrease in g_m since channel resistance would become larger. For a given gate length, L, the "V_{ds} = 15 V" curve is observed to have a higher f_T value than the others for the same reason as in Fig. 4.16 – a higher V_{ds} value decreases the C_{gd} and C_{gs} values. For a gate length of 1μm, drain biases of 15 V, 10 V, and 5 V yielded f_T values of 13.15 GHz, 12.69 GHz, and 11.71 GHz, respectively. Decreasing the gate length by one order of magnitude to 0.1 μm drastically increased the cutoff frequency of the device for all drain biases to values of 199.6 GHz (V_{ds} = 15 V), 183.34 GHz (V_{ds} = 10 V), and 147.64 GHz (V_{ds} = 5 V). Rashmi *et al.* [23] and Ambacher *et al.* have reported similar findings regarding the cut-off frequency and its relationship with the gate length. The results shown in [10,40] are quite close to what has been shown here, revealing the model's ability to predict device cut-off frequency values.

Figures 4.16 and 4.17 have shown that the frequency response of a GaN HEMT is most easily improved by either increasing the drain voltage or decreasing the gate length. This is because both methods reduce the magnitude of the gate-to-source and gate-to-drain capacitances. Although the process for fabricating

GaN HEMT devices has not yet achieved a mature stage, these results certainly provide a good reason for continuing research in the area of HEMT modeling for high-power high-frequency circuits.

Chapter 5

Conclusions and Future Research

5.1 Conclusion

A physics-based model has been proposed to model High Electron Mobility Transistors. The Schrödinger equation was used in conjunction with Fermi-Dirac statistics to generate an equation relating the sheet carrier density, n_s, to the Fermi level, E_f. Two quantum well eigenstates were considered, limiting the effectiveness of the model in predicting n_s concentrations to V_{gs} values below threshold. A simplified version of the Poisson equation was used to obtain a linear charge-control relationship between n_s and V_{gs} which was later used to derive I/V equations. The linear and saturation I/V equations included both current components which was found to improve upon previous results reported by other authors. Discrepancies between the calculated and experimentally measured transconductance values supports the argument that more research and work needs to be done to improve the accuracy of performance-predicting ability of the model. Most of the other calculated results lacked experimental data to be compared against, in which case their profiles and magnitudes were validated by comparison to experimental GaAs data. This was the case for most of the conductance calculations and all of the capacitance and frequency calculations.

Most all the results were very encouraging in that all fell within 22% of the experimentally measured values (when measured values were available). In reality, the agreement between simulated and measured results was much better than that for the I_{ds} vs. V_{ds} and I_{ds} vs. V_{gs} curves where the percent deviation between the results never became larger than 10% for any curve at any bias condition. Percent errors could not be calculated for the other curves since they were not able to be explicitly compared with experimental measurements.

The most notable weakness in the model is its inability to accurately predict transconductance values in and around the transitional area between the saturation and linear regions. This is a serious limitation considering that peak transconductance occurs when the device transitions from the saturation region to the

linear region. Since g_m is derived from the current curves, an explanation may be found there. When looking at Fig. 4.8, good agreement was found between predicted and measured results with only a slight bump occurring in the curves where the saturation and linear regions meet. The largest differences between the measured and predicted values came either in the linear region or when thermal effects degraded the current performance of the device. These differences, although small in the current curves, become magnified and more pronounced with respect to the conductance parameters (g_m and g_0). Unfortunately, experimental data could not be obtained for profile comparison for output conductance, g_0. References that were found provided simulation results that closely matched the results presented in Chapter 4. This was the case with the remainder of results that were presented. They could not be compared to real samples so their accuracy could not be validated. It is worthwhile to mention that excellent agreement was found between Figures 4.1, 4.12, 4.15, 4.16, and 4.17 and their respective references. Figures 4.3 and 4.14, however, did not have the same "success." The discrepancy with these curves involved their profiles more so than their magnitudes. The n_s vs. V_{gs} relationship pictured in Fig. 4.3 shows a near-perfect linear relationship, with n_s having negligible magnitudes for V_{gs} values less than threshold. In actuality, when V_{gs} decreases to near threshold, n_s begins to decrease asymptotically with respect to the x-axis [15] (much like in Fig. 4.1 where n_s is plotted against E_f). In Fig, 4.14, C_{gs} is shown to increase at a decreasing rate for the V_{gs} range specified. Although this is generally what happens, C_{gs} should rise faster and stabilize in value with increasing V_{gs} [36].

5.2 Future Research

The following list of statements has been offered to detail the weaknesses of the model and of the GaN growth process and to provide reasoning for future research:

1) Although major progress has been achieved in recent years, the GaN growth process still needs work. Steady improvement has been achieved in decreasing contact resistance values, idealizing junction profiles, and increasing the quality of the grown semiconductor. More progress is needed, however, until the process can produce acceptable yield. The rate at which the process can be refined is limited by the fact that there exist only a handful of fabrication facilities worldwide that

the following V_{th} discrepancy was found by the author. In Garrido *et al.* [16], V_{th} values of -7 V and -4.2 V were reported for two samples, one with an m of 25% and d_d of 350 Å, and the other with an m of 22% and d_d of 280 Å, respectively. Both samples had an AlGaN spacer layer thickness of 30 Å and Si doping density of $2x10^{18}$ cm^{-3} in the charge donor layer. Using the information provided in [16] and equation (2.6), V_{th} values of -10.14 V and -6.72 V were calculated, in clear disagreement with the measured values. The difference between calculated V_{th} and measured V_{th} would lead directly to discrepancies with all other results, nullifying the accuracy of the model. This discovery had a large part in deciding which hardware results would be used for comparison and which could not be used. Using equation (2.6) and the device parameters reported in Wu *et al.* [23], good agreement was found between the calculated V_{th} value and the V_{th} reported (~ -2.8 V) which was the motivation behind using [23] for experimental comparison. As measured results would suggest, the device threshold voltage does not vary with mole fraction as equation (2.6) and Fig. 5.1 would suggest. With that said, future research modeling their relationship would be of great value in that it would not limit the applicability of the model to devices with mole fractions of 15% or less.

3) The proposed model focused only on the strong inversion region, leading to poor subthreshold n_s concentration predictions. This is because Schrödinger's equation was used to obtain only two quantum well eigenstates. Certainly if more eigenstates were included, and the Poisson equation was not simplified, excellent agreement could be expected between predicted and experimentally measured n_s values in the subthreshold region. Nonetheless the V_{th} vs. m problem would still need to be resolved if agreement between n_s and V_{gs} was desired for devices of mole fractions higher than 15%. It was unnecessary for the proposed model to accurately predict n_s values in the subthreshold region since the GaN-based HEMT is being considered as a high-power amplification device.

4) In Fig. 4.7, the experimental data plotted increased, peaked, and then decreased slightly as the drain bias continued to increase. Thermal effects were offered as the explanation for this behavior. Clearly, if a device heats up, the ability of the carriers to move (mobility) in the presence of an

electric field will decrease. This affects the channel current in a negative manner which is the reasoning behind the decrease in current observed at elevated drain biases. Since the proposed model did not take these effects into account, no associated decrease was observed in the calculated results, therefore increasing the deviation between experimental and calculated values. Attempts have been made to alleviate this problem by growing GaN circuits on top of SiC substrates that have high thermal conductivity (around 4-5 $W \cdot cm^{-1}$ as opposed to 0.35 $W \cdot cm^{-1}$ for sapphire substrates [41]). Although thermal effects still cause device performance to degrade at high drain biases (V_{ds} > 20V), the degradation is less pronounced compared to when sapphire substrates are used. A future model that includes performance-degrading effects such as thermal effects and trapping effects would be useful in that it could predict more accurately the performance of the device.

The model improvements just listed, with serve as motivation for future research in this area. Modeling the relationship between the threshold voltage and mole fraction is most important since equation (2.6) limits the proposed model to devices of mole fraction 15% or less.

References

[1] William Liu, *Fundamentals of III-V Devices: HBTs, MESFETs, and HFETs/HEMTs*, John Wiley and Sons, New York, NY, 1999.

[2] Fujitsu Quantum Devices, LTD. – History of HEMT Transistors. Retrieved April 26, 2004 from http://eesof.tm.agilent.com/docs/iccap2002/MDLGBOOK/7DEVICE_MODELING/3TRANSISTORS/0History/HEMTHistory.pdf.

[3] Y. Zhang, I.P. Smorchkova, C.R. Elsass, S. Keller, J.P. Ibbetson, S. Denbaars, U.K. Mishra, "Charge control and mobility in AlGaN/GaN transistors: Experimental and theoretical studies," *Journal of Applied Physics*, vol. 87, no. 11, pp. , June, 2000.

[4] Michael E. Levinshtein, Sergey L. Rumyantsev, Michael S. Shur, *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe*, Wiley-Interscience, New York, NY, 2001.

[5] Ioffe Physico-Technical Institute – Physical Properties of Semiconductors. Retrieved April 29, 2004 from http://www.ioffe.ru/SVA/NSM/Semicond/index.html.

[6] O. Ambacher, J. Smart, J.R. Shealy, N.G. Weimann, K. Chu, M. Murphy, W.J. Schaff, L.F. Eastman, "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures," *Journal of Applied Physics*, vol. 85, no. 6, pp. 3222-3233, March, 1999.

[7] R. Thompson, T. Prunty, V. Kaper, J.R. Shealy, "Performance of the AlGaN HEMT Structure With a Gate Extension," *IEEE Trans. on Electron Devices*, vol. 51, no. 2, pp. 292-295,February, 2004.

[8] W. Saito, Y. Takada, M. Kuraguchi, K. Tsuda, I. Omura, T. Ogura, H. Ohashi, "High Breakdown Voltage AlGaN-GaN Power-HEMT Design and High Current Density Switching Behavior," *IEEE Trans. on Electron Devices*, vol. 50, no. 12, pp. 2528-2531, December, 2003.

[9] W. Lu, V. Kumar, E.L. Piner, I. Adesida, "DC, RF, and Microwave Noise Performance of AlGaN-GaN Field Effect Transistors Dependence of Aluminum Concentration," *IEEE Trans. on Electron Devices*, vol. 50, no. 4, pp. 1069-1074, April, 2003.

[10] S. Binari, K. Ikossi, J. Roussos, W. Kruppa, D. Park, H.B. Dietrich, D.D. Koleske, A.E. Wickenden, R.L. Henry, "Trapping Effects and Microwave Power Performance in AlGaN/GaN HEMTs," *IEEE Trans. on Electron Devices*, vol. 48, no. 3, pp. 465-471, March, 2001.

[11] P.M. Asbeck, E.T. Yu, S.S. Lau, G.J. Sullivan, J.V. Hove, J. Redwing, "Piezoelectric charge densities in AlGaN/GaN HFETs," *Electronic Letters*, vol. 33, no. 14, pp. 1230-1231, July, 1997.

[12] Michael Shur, *GaAs devices and circuits*, Plenum, New York, NY, 1987.

[13] Patrick Roblin and Hans Rohdin, *High-speed heterostructure devices*, Cambridge University Press, Cambridge, UK, 2002.

[14] A. Shey, W.H. Ku, "An Analytical Current-Voltage Characteristics Model for High Electron Mobility Transistors Based on Nonlinear Charge-Control Formulation," *IEEE. Trans. on Electron Devices*, vol. 36, no. 10, pp. 2299-2309, October, 1989.

[15] Rashmi, A. Kranti, S. Haldar, R.S. Gupta, "An accurate charge control model for spontaneous and piezoelectric polarization dependent two-dimensional electron gas sheet charge density of lattice-mismatched AlGaN/GaN HEMTs," *Solid-State Electronics*, vol. 46, pp. 621-630, November, 2002.

[16] J.A. Garrido, J.L. Sanchez-Rojas, A. Jimenez, E. Munoz, F. Omnes, P. Gibart, "Polarization fields determination in AlGaN/GaN heterostructure field-effect transistors from charge control analysis," *Applied Physics Letters*, vol. 75, no. 16, pp. 2407-2409, October, 1999.

[17] Rashmi, A. Kranti, S. Haldar, M. Gupta, R.S. Gupta, "Comprehensive Analysis of Small-Signal Parameters of Fully Strained and Partially Relaxed High Al-Content Lattice Mismatched $Al_mGa_{1-m}N$/GaN HEMTs," *IEEE Trans. on Microwave Theory and Techniques*, vol. 51, no. 2, pp. 607-617, February, 2003.

[18] Rashmi, S. Haldar, R.S. Gupta, "2-D Analytical Model for Current-Voltage Characteristics and Output Conductance of AlGaN/GaN MODFET," *Microwave and Optical Tech. Letters*, vol. 29, no. 2, pp. 117-123, April, 2001.

[19] P.P. Ruden, J.D. Albrecht, A. Sutandi, S.C. Binari, K. Ikossi-Anastasiou, M.G. Ancona, R.L. Henry, D.D. Koleske, A.E. Wickenden, "Extrinsic Performance Limitations of AlGaN/GaN Heterostructure Field Effect Transistors," *MRS Internet J. Nitride Semicond. Res 4S1*, G6.35 (1999).

[20] O. Ambacher, B. Foutz, J. Smart, J.R. Shealy, N.G. Weimann, K. Chu, M. Murphy, A.J. Sierakowski, W.J. Schaff, L.F. Eastman, "Two dimensional electron gases induces by spontaneous and piezoelectric polarization in undoped and doped AlGaN/GaN heterostructures," *J. of Applied Physics*, vol. 87, no. 1, pp. 334-344, January, 2000.

[21] V. Fiorentini, F. Bernardini, O. Ambacher, "Evidence for nonlinear macroscopic polarization in III-V nitride alloy heterostructures," *Applied Physics Letters*, vol. 80, no. 7, pp. 1204-1206, February, 2002.

[22] C.S. Chang, H.R. Fetterman, "An Analytic Model for HEMT's Using New Velocity-Field Dependence," *IEEE Trans. on Electron Devices*, vol. ED-34, no. 7, pp. 1456-1462, July, 1987.

[23] Y.F. Wu, S. Keller, P. Kozodoy, B.P. Keller, P. Parikh, D. Kapolnek, S.P. Denbaars, U.K. Mishra, "Bias Dependent Microwave Performance of AlGaN/GaN MODFET's Up To 100 V," *IEEE Electron Device Letters*, vol. 18, no. 6, pp. 290-292, June, 1997.

[24] E.J. Miller, X.Z. Dang, E.T. Yu, "Gate leakage current mechanisms in AlGaN/GaN heterostructure field-effect transistors," *J. of Applied Physics*, vol. 88, no. 10, pp. 5951-5958, November, 2000.

[25] E.T. Yu, X.Z. Dang, L.S. Yu, D. Qiao, P.M. Asbeck, S.S. Lau, G.J. Sullivan, K.S. Boutros, J.M. Redwing, "Schottky barrier engineering in III-V nitridesvia the piezoelectric effect," *Applied Physics Letters*, vol. 73, no. 13, pp. 1880-1882, September, 1998.

[26] David Halliday, Robert Resnick, and Jearl Walker, *Fundamentals of Physics: Fifth Edition*, John Wiley and Sons, New York, NY, 1997.

[27] Y. Zhang, J. Singh, "Charge control and mobility studies for an AlGaN/GaN high electron mobility transistor," *J. of Applied Physics*, vol. 85, no. 1, pp. 587-594, January, 1999.

[28] Y. Ando, T. Itoh, "Analysis of Charge Control in Pseudomorphic Two-Dimensional Electron Gas Field-Effect Transistors," *IEEE Trans. on Electron Devices*, vol. 35, no. 12, pp. 2295-2301, December, 1988.

[29] L. Eastman, V. Tilak, J. Smart, B.M. Green, E.M. Chumbes, R. Dimitrov, H. Kim, O.S. Ambacher, N. Weimann, T. Prunty, M. Murphy, W.J. Schaff, J.R. Shealy, "Undoped AlGaN/GaN HEMTs for Microwave Power Amplification," *IEEE Trans. on Electron Devices*, vol. 48, no. 3, pp. 479-485, March, 2001.

[30] M. Wolter, P. Javorka, A. Fox, M. Marso, H. Luth, P. Kordos, R. Carius, A. Alam, M. Heuken, "Photo-Ionization Spectroscopy of Traps in AlGaN/GaN High Electron Mobility Transistors," *J. of Electronic Materials*, vol. 31, no. 12, pp. 1321-1324, December, 2002.

[31] X. Hu, G. Simin, J. Yang, M.A. Khan, R. Gaska, M.S. Shur, "Enhancement mode AlGaN/GaN HFET with selectively grown *pn* junction gate," *Electronic Letters*, vol. 36, no. 8, pp. 753-754, April, 2000.

[32] A. Mahajan, M. Arafa, P. Fay, C. Caneau, I. Adesida, "Enhancement-Mode High Electron Mobility Transistors (E-HEMT's) Lattice-Matched to InP," *IEEE Trans. on Electron Devices*, vol. 45, no. 12, pp. 2422-2429, December, 1998.

[33] T. Egawa, H. Ishikawa, T. Jimbo, M. Umeno, "Recessed Gate AlGaN/GaN MODFET on Sapphire Grown by MOCVD," *IEDM Technical Digest* 1999:401-4.

[34] Rashmi, A. Agrawal, S. Sen, S. Haldar, R.S. Gupta, "Analytical Model for DC Characteristics and small-signal parameters of AlGaN/GaN Modulation-Doped Field-Effect Transistor for Microwave Circuit Applications," *Microwave and Optical Technology Letters*, vol. 27, no. 6, pp. 413-419, December, 2000.

[35] A. Fox, M. Marso, P. Javorka, P. Kordos, "RF Small-Signal and Power Characterization of AlGaN/GaN HEMTs," *ASDAM Technical Digest* 2002:291-294.

[36] J.W. Lee, K.J. Webb, "A Temperature-Dependent Nonlinear Analytic Model for AlGaN-GaN HEMTs on SiC," *IEEE Trans. on Microwave Theory and Techniques*, vol. 52, no. 1, pp. 2-7, January, 2004.

[37] J. Michael Golio, *Microwave MESFETs & HEMTs*, Artech House, Boston, MA, 1991.

[38] U.K. Mishra, Y.F. Wu, B.P. Keller, S. Keller, S.P. Denbaars, "GaN based microwave power HEMTs," *Proc. Int. Physics of Semiconductor Devices Workshop*, 1998, pp. 878-883.

[39] W. Lu, J. Yang, M.A. Khan, I. Adesida, "AlGaN/GaN HEMTs on SiC with over 100 GHz f_T and Low Microwave Noise," *IEEE Trans. on Electron Devices*, vol. 48, no. 3, pp. 581-585, March, 2001.

[40] F. Schwierz, O. Ambacher, "Recent Advances in GaN HEMT Development," *The 11th IEEE Int. Symp. on Electron Devices for Microwave and Optoelectronic Applications*, pp. 204-209, November, 2003.

[41] R. Gaska, Q. Chen, J. Yang, A. Osinsky, M.A. Khan, M.S. Shur, "High-Temperature Performance of AlGaN/GaN HFET's on SiC Substrates" *IEEE Electron Device Letters*, vol. 18, no. 10, pp. 492-494, October, 1997.

Appendix I – $I_{d,lin}$ Derivation

Equation (2.12) shows the current density equation after plugging in the mobility model (equation (2.10)),

$$I_{ds}\left(1+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)\frac{dV_c(x)}{dx}\right)=Wq\mu_0\left(n_s(m,x)\frac{dV_c(x)}{dx}+\frac{k_B T}{q}\frac{dn_s(m,x)}{dx}\right) \tag{2.12}$$

Substituting equation (2.11) for $n_s(m,x)$ yields,

$$I_{ds}\left(1+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)\frac{dV_c(x)}{dx}\right)=Wq\mu_0\left(\frac{\varepsilon(m)}{qd}\left(V_{gt}-V_c(x)\right)\frac{dV_c(x)}{dx}+\frac{k_B T}{q}\frac{d}{dx}\left(\frac{\varepsilon(m)}{qd}\left(V_{gt}-V_c(x)\right)\right)\right) \tag{2.13a}$$

where

$$d = d_d + d_i + \Delta d$$

$$V_{gt} = V_{gs} - V_{th}(m)$$

Distributing $\frac{d}{dx}$ through on the right hand side (RHS) of (2.13a),

$$I_{ds}\left(1+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)\frac{dV_c(x)}{dx}\right)=Wq\mu_0\left(\frac{\varepsilon(m)}{qd}\left(V_{gt}-V_c(x)\right)\frac{dV_c(x)}{dx}-\frac{k_B T}{q}\left(\frac{\varepsilon(m)}{qd}\frac{dV_c(x)}{dx}\right)\right) \tag{2.13b}$$

Simplifying the RHS and multiplying both sides by dx yields,

$$I_{ds}\left(dx+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)dV_c(x)\right)=\frac{W\mu_0\varepsilon(m)}{d}\left(V_{gt}-V_c(x)-\frac{k_B T}{q}\right)dV_c(x) \tag{2.13c}$$

The boundary conditions represent the intrinsic voltages at either end of the channel and are listed here for reference:

$$V_c(x)\big|_{x=0} = I_{ds}R_s$$

$$V_c(x)\big|_{x=L} = V_{ds} - I_{ds}R_d$$

Since the intrinsic voltages are known at either end of the channel, (2.13c) can be integrated from $x = 0$ to $x = L$ to obtain an expression for the drain current in the linear region,

$$I_{ds}\left(\int_0^L dx + a\int_0^L dV_c(x)\right) = b\left(V_{gt} - \frac{k_B T}{q}\right)\int_0^L dV_c(x) - b\int_0^L V_c(x)dV_c(x) \qquad (2.13d)$$

where

$$a = \frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}$$

$$b = \frac{W\mu_0\varepsilon(m)}{d}$$

Performing the integration yields,

$$I_{ds}\left(L + aV_c(x)\big|_{x=0}^{x=L}\right) = b\left(V_{gt} - \frac{k_B T}{q}\right)V_c(x)\big|_{x=0}^{x=L} - \frac{b}{2}V_c^2(x)\big|_{x=0}^{x=L} \qquad (2.13e)$$

Plugging in the boundary conditions and simplifying,

$$I_{ds}L + I_{ds}V_{ds}a - aI_{ds}^2(R_d + R_s) = bV_{gs}^{'}V_{ds} - bV_{gs}^{'}I_{ds}(R_d + R_s) - \frac{b}{2}\left(V_{ds}^2 - 2V_{ds}I_{ds}R_d + I_{ds}^2\left(R_d^2 - R_s^2\right)\right) \qquad (2.13f)$$

where

$$V_{gs}^{'} = V_{gt} - \frac{k_B T}{q}$$

By combining the coefficients of the I_{ds} terms, equation (2.13f) can be prepared for the quadratic formula,

$$0 = \alpha I_{ds}^2 + \beta I_{ds} + \gamma \qquad (2.13g)$$

and

$$I_{d.lin} = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} \qquad (2.13)$$

where

$$\alpha = a(R_d + R_s) - \frac{b}{2}\left(R_d^2 - R_s^2\right)$$

$$\beta = bV_{ds}R_d - bV_{gs}^{'}(R_d + R_s) - L - aV_{ds}$$

$$\gamma = bV_{gs}^{'}V_{ds} - \frac{b}{2}V_{ds}^2$$

Appendix II – $V_c(x)$ Derivation

This appendix outlines the derivation of the channel potential, $V_c(x)$. Obtaining this expression was pertinent to the model's incorporation of channel length modulation occurring in the saturation region. The expression obtained here is later used to form the bias-dependent expression for L_l, the length of the Low-Field region in the channel.

Starting with equation (2.9) and plugging in equations (2.10) and (2.11) yields,

$$I_{ds}\left(1+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)\frac{dV_c(x)}{dx}\right)=Wq\mu_0\left(\frac{\varepsilon(m)}{qd}\left(V_{gt}-V_c(x)\right)\frac{dV_c(x)}{dx}+\frac{k_BT}{q}\frac{d}{dx}\left(\frac{\varepsilon(m)}{qd}\left(V_{gt}-V_c(x)\right)\right)\right) \tag{2.14a}$$

Simplifying (2.14a) gives the same expression as (2.13c) and is repeated here for reference,

$$I_{ds}\left(dx+\left(\frac{\mu_0 E_c - v_{sat}}{E_c v_{sat}}\right)dV_c(x)\right)=\frac{W\mu_0\varepsilon(m)}{d}\left(V_{gt}-V_c(x)-\frac{k_BT}{q}\right)dV_c(x) \tag{2.14b}$$

Since the channel is pinched off and is of unknown length, only the boundary condition corresponding to $x = 0$ can be used. Integrating (2.14b) from 0 to x gives,

$$I_{ds}\left(\int_0^x dx + a\int_0^x dV_c(x)\right)=b\left(V_{gt}-\frac{k_BT}{q}\right)\int_0^x dV_c(x)-b\int_0^x V_c(x)dV_c(x) \tag{2.14c}$$

After integrating,

$$I_{ds}\left(x+aV_c(x)\Big|_{x=0}^{x=x}\right)=b\left(V_{gt}-\frac{k_BT}{q}\right)V_c(x)\Big|_{x=0}^{x=x}-\frac{b}{2}V_c^2(x)\Big|_{x=0}^{x=x} \tag{2.14d}$$

Plugging in the boundary conditions yields,

$$I_{ds}\left(x+a\left(V_c(x)-I_{ds}R_s\right)\right)=bV_{gs}'\left(V_c(x)-I_{ds}R_s\right)-\frac{b}{2}\left(V_c^2(x)-I_{ds}^2R_s^2\right) \tag{2.14e}$$

Distributing the coefficients through and combining terms with like orders of $V_c(x)$ simplifies (2.14e) to,

$$\frac{b}{2}V_c^2(x)+\left(aI_{ds}-bV_{gs}^{'}\right)V_c(x)+\left(xI_{ds}+bV_{gs}^{'}I_{ds}R_s-aI_{ds}^2R_s-\frac{b}{2}I_{ds}^2R_s^2\right)=0 \qquad (2.14)$$

The quadratic formula is used to obtain a solution to (2.14),

$$V_c(x)=\frac{-\beta-\sqrt{\beta^2-4\alpha\gamma}}{2\alpha}$$

where

$$\alpha=\frac{b}{2}$$

$$\beta=aI_{ds}-bV_{gs}^{'}$$

$$\gamma=xI_{ds}+bV_{gs}^{'}I_{ds}R_s-aI_{ds}^2R_s-\frac{b}{2}I_{ds}^2R_s^2$$